RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Novel physico-chemical mechanism of the mutagenic tautomerisation of the Watson-Crick-like A·G and C·T DNA base mispairs: a quantum-chemical picture

Ol'ha O. Brovarets'[a,b] & Dmytro M. Hovorun[a,b,✉]

[a]Department of Molecular and Quantum Biophysics, Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, 150 Akademika Zabolotnoho Str., 03680 Kyiv, Ukraine

[b]Department of Molecular Biotechnology and Bioinformatics, Institute of High Technologies, Taras Shevchenko National University of Kyiv, 2-h Akademika Hlushkova Ave., 03022 Kyiv, Ukraine

✉Corresponding author. E-mail: dhovorun@imbg.org.ua

Abstract. Newly discovered physico-chemical mechanism of the mutagenic tautomerisation of the long A·G and short C·T Watson-Crick DNA base mispairs was revealed for the first time. Tautomerisation of each mismatch occurs *via* four topologically and energetically different ways through highly stable transition states – H-bonded tight ion pairs containing protonated and deprotonated bases. These processes are accompanied by a significant rebuilding of the base mispairs with Watson-Crick architecture into the mismatches wobbled towards both minor and major DNA grooves and *vice versa*. Moreover, it was established that these tautomerisation reactions occur non-dissociatively and are accompanied by the consequent replacement of the unique patterns of the intermolecular specific interactions along IRC. Finally, we briefly discuss the possible biological significance of the obtained results for clarifying the microstructural foundations of the origin of the spontaneous point mutations within the framework of the classical Watson-Crick tautomeric hypothesis.

Keywords: DNA biosynthesis · Spontaneous point mutation · Replication error · Watson-Crick-like mispair · Wobble mismatch · DPT · Mutagenic tautomer · MP2 and B3LYP · QTAIM

† Electronic supplementary information (ESI) available: (i) Energetic characteristics of the investigated DNA mispairs; (ii) Profiles of the glycosidic parameters along the IRC of tautomerisation; (iii) Energetic profiles and parameters of the mismatches planarization. See DOI: 10.1039/x0xx00000x.

Introduction.

A special place among all incorrect DNA base pairs – the "culprits" of the occurrence of the spontaneous point mutations [1-4] – is occupied by two base pairs – the so-called long A·G and short C·T [5-12] pairs with Watson-Crick molecular architecture of the binding between bases [13,14]. The DNA bases, that form the A·G and C·T DNA base mispairs, are in the main tautomeric form [15-17] in contrast to the other mispairs. Notably, extension efficiencies for the mismatched base pairs presented by Goodman M. *et al.* [18] constitute $10^{-4}\div10^{-5}/10^{-2}$ for the C(template)·T(primer)/T·C·mispairs, respectively, while for the G·A/A·G mismatches they are less than 10^{-6} comparably to the canonical A·T Watson-Crick DNA base pair.

We have previously investigated in details the process of the DPT tautomerisation of these pairs as their intrinsic property [11,12], which is important in terms of the mutations fixation in the subsequent rounds of DNA replication. As a result, we have concluded that dissociation of the Watson-Crick-like A·G [11] and C·T [12] mispairs into the isolated monomers by DNA replication machinery proceeds without changing of their tautomeric status. However, this conclusion is valid if and only if the DPT tautomerisation of the aforementioned pairs occurs *via* the classical Löwdin's mechanism [19,20].

In this paper, we describe for the first time new mechanism of the mutagenic tautomerisation, combined by proton transfer and shifting of the bases relatively to each other, of the long A·G and short C·T DNA base mispairs with Watson-Crick (WC) architecture. During this tautomeric conversion the Watson-Crick-like base pairs undergo large-scale structural changes to adopt wobble (w) geometry. Characteristic difference of this novel mechanism from the Löwdin's mechanism consists in the fact that the mutagenic tautomerisation of the pairs is accompanied by a significant changes in their geometry, namely by the transition into the wobble configuration, and is carried out through the highly stable transition states (TS) [21,22] $A^{+}\cdot G^{-}$, $A^{-}\cdot G^{+}$, $C^{+}\cdot T^{-}$ and $C^{-}\cdot T^{+}$ (signs "+" and "–" denote protonated and deprotonated DNA bases, respectively). Obtained data shed light on the nature of the spontaneous point A·G/G·A and C·T/T·C replication errors in DNA, when for one reason or another [13,14,19,20] complementary DNA bases randomly change their canonical tautomeric status into the mutagenic during DNA replication. Results presented in this paper

can broaden our outlook and provide valuable insights into the mechanisms of the origin of the spontaneous point replication errors at the atomistic level.

Computational Methods.

All calculations of the geometries and harmonic vibrational frequencies of the considered base mispairs and transition states of their conversion have been performed using Gaussian'09 package [23] at the B3LYP DFT/6-311++G(d,p) level of theory [24-26], that has been applied for analogous systems and verified to give accurate geometrical structures, normal mode frequencies, barrier heights and characteristics of intermolecular H-bonds [27,28]. A scaling factor that is equal to 0.9668 has been applied in the present work for the correction of the harmonic frequencies of all studied base pairs [29,30]. We have confirmed the minima and TS, located by means of Synchronous Transit-guided Quasi-Newton method [31,32], on the potential energy landscape by the absence or presence, respectively, of the imaginary frequency in the vibrational spectra of the complex.

In order to consider electronic correlation effects as accurately as possible, we followed geometry optimizations with single point energy calculations using MP2 functional [33] and a wide variety of basis sets, in particular, Pople's basis sets of valence triple-ζ quality [34,35], as well as Dunning's cc-type basis sets [36], augmented with polarization and/or diffuse functions: 6-311++G(2df,pd), 6-311++G(3df,2pd), cc-pVTZ and cc-pVQZ.

Reaction pathways have been established by following intrinsic reaction coordinate (IRC) in the forward and reverse directions from each TS using Hessian-based predictor-corrector integration algorithm [37,38] with tight convergence criteria. These calculations eventually ensure that the proper reaction pathway, connecting the expected reactants and products on each side of the TS, has been found. We've have investigated the evolution of the energetic and geometric characteristics of the H-bonds and base pairs along the reaction pathway establishing them at each point of the IRC [28,29].

Electronic interaction energies E_{int} have been calculated at the MP2/6-311++G(2df,pd) level of theory as the difference between the total energy of the base mispair and the energies of the isolated monomers. Gibbs free energy of interaction has been obtained using similar equation. In each case the interaction energy was corrected for the basis set superposition error (BSSE) [39,40] through the counterpoise procedure [41,42].

The Gibbs free energy G for all structures was obtained in the following way:

$$G=E_{el}+E_{corr}, \quad (1)$$

where E_{el} – electronic energy, while E_{corr} – thermal correction. We applied the standard TS theory [43] to estimate the activation barriers of the tautomerisation reaction.

The time $\tau_{99.9\%}$ necessary to reach 99.9% of the equilibrium concentration of the reactant and product in the system of reversible first-order forward (k_f) and reverse (k_r) reactions can be estimated by formula [43]:

$$\tau_{99.9\%} = \frac{ln10^3}{k_f + k_r}. \quad (2)$$

To estimate the values of the rate constants k_f and k_r:

$$k_{f,r} = \Gamma \cdot \frac{k_B T}{h} e^{-\frac{\Delta\Delta G_{f,r}}{RT}} \quad (3)$$

we applied standard TS theory [43], in which quantum tunneling effect are accounted by Wigner's tunneling correction [44], that has been successfully used for the DPT reactions [45-47]:

$$\Gamma = 1 + \frac{1}{24}\left(\frac{h\nu_i}{k_B T}\right)^2, \quad (4)$$

where k_B – Boltzmann's constant, h – Planck's constant, $\Delta\Delta G_{f,r}$ – Gibbs free energy of activation for the tautomerisation reaction in the forward (f) and reverse (r) directions, ν_i – magnitude of the imaginary frequency associated with the vibrational mode at the TSs.

Bader's quantum theory of Atoms in Molecules (QTAIM) was applied to analyse the electron density distribution [48]. The topology of the electron density was analysed using AIMAll program package [49] with all default options. The presence of a bond critical point (BCP), namely the so-called (3,-1) BCP, and a bond path between hydrogen donor and acceptor, as well as the positive value of the Laplacian at this BCP ($\Delta\rho>0$), were considered as criteria for the H-bond formation [50,51]. Wave functions were obtained at the level of theory used for geometry optimisation.

The energies of the N2H···HC2 dihydrogen (DH) bond [11,52] in the A·G(WC) base mispair, the $N2^{+}H\cdots HC2^{-}$ DH-bond in the $TS^{A-\cdot G+}_{A\cdot G(WC)\leftrightarrow A^{*}\cdot G\downarrow(w)}$ transition state, the weak C2H···O6 H-bond in the $A^{*}\cdot G_{\downarrow}(w)$ mismatch, the attractive [12] O2···O2 van der Waals contacts in the C·T(WC), C*·T*(WC) base mispairs and $TS_{C\cdot T(WC)\leftrightarrow C^{*}\cdot T^{*}(WC)}$ transition state, and

also the attractive [12] $O2^{+}\cdots N3^{-}$ van der Waals contact in the $TS^{C-\cdot T+}{}_{C\cdot T(WC)\leftrightarrow C\cdot T^{*}\uparrow(w)}$ transition state were calculated by the empirical Espinosa-Molins-Lecomte (EML) formula [53-54] based on the electron density distribution at the (3,-1) BCPs of the H-bonds:

$$E_{NH\cdots HC/N}{}^{+}{}_{H\cdots HC}{}^{-}{}_{/CH\cdots O/O\cdots O/O}{}^{+}{}_{\cdots N}{}^{-}=0.5\cdot V(r), \quad (5)$$

where V(r) – value of a local potential energy at the (3,-1) BCP.

The energies of the N4H···N3 and N3H···N3 H-bonds in the $TS_{C^{*}\cdot T\uparrow(w)\leftrightarrow C\cdot T^{*}O2(w)}$ and $TS_{C\cdot T(WC)\leftrightarrow C^{*}\cdot T^{*}(WC)}$ transition states, respectively, containing loosened covalent bridges were estimated by the Nikolaienko-Bulavin-Hovorun formula [55]:

$$E_{NH\cdots N}=-2.03+225\cdot\rho, \quad (6)$$

where ρ – the electron density at the (3,-1) BCP of the H-bond.

The energies of all others AH···B conventional H-bonds were evaluated by the empirical Iogansen's formula [56]:

$$E_{AH\cdots B}=0.33\cdot\sqrt{\Delta\nu-40}, \quad (7)$$

where Δν – magnitude of the frequency shift of the stretching mode of the AH H-bonded group involved in the AH···B H-bond relatively the unbound group. The partial deuteration was applied to minimize the effect of vibrational resonances [45,46].

One and the same values of the ν frequency and d_{AH} distances shifts have been observed for the different H-bonds in the $TS^{C-\cdot T+}{}_{C\cdot T(WC)\leftrightarrow C^{*}\cdot T\downarrow(w)}$ and $TS^{C+\cdot T-}{}_{C^{*}\cdot T^{*}(WC)\leftrightarrow C^{*}\cdot T\uparrow(w)}$ transition states, since they involve joint $O4^{+}H/N3^{+}H/N4^{+}H$ donor groups. The same data have been observed for the various H-bonds in the $TS^{A+\cdot G-}{}_{A\cdot G(WC)\leftrightarrow A^{*}\cdot G\uparrow(w)}$ and $TS^{A+\cdot G-}{}_{A\cdot G(WC)\rightarrow A\cdot G^{*}\downarrow(w)}$ transition states containing mutual $N6^{+}H/N1^{+}H$ donor groups.

The atomic numbering scheme for the DNA bases is conventional [57].

Results and their discussion.

First of all, before proceeding to the presentation of the obtained results and their discussion, we would like to generally outline ideas, which have encouraged us for the discovery of the new mechanism of the mutagenic tautomerisation *via* the sequential DPT accompanied with structural rearrangements of the DNA bases within the A·G and C·T base mispairs with Watson-Crick type of the H-bonding. It is known for certain that canonical DNA bases – adenine (A), guanine (G), cytosine (C) and thymine (T) – are able, in principle, to

transfer from the canonical into the mutagenic tautomeric form by intramolecular migration of the amino proton of the A and C bases to the neighboring N1 or N3 nitrogen atoms, respectively, or by intramolecular migration of the imino proton of the G and T bases to the neighboring O6 and O4 oxygen atoms, respectively [58-64]. However, these tautomeric transitions occur quite slowly ($\sim 10^{10} \div 10^{20}$ s), even in comparison with the time of the DNA replication in the cell ($\sim 10^{6}$ s [65,66]), since very high activation energy barriers (32÷46 kcal·mol^{-1}) correspond to such conversions [58,60,63,64]. We have hypothesized [21,22] that each of the canonical DNA bases in the considered A·G and C·T DNA base mispairs can catalyze mutagenic tautomerisation of the "complementary" base, firstly nipping off the migrating proton, then shifting relatively other base within the base pair into the major or minor groove sides of the DNA helix and adding this mobile or other acidic proton to the neighboring nitrogen or oxygen atoms of the "complementary" base. It is quite natural to expect that transition states of these processes would represent itself highly stable structures [21,22], namely $A^{+}\cdot G^{-}$, $A^{-}\cdot G^{+}$, $C^{+}\cdot T^{-}$ and $C^{-}\cdot T^{+}$ H-bonded tight ion pairs, geometry of which is no longer Watson-Crick, but is not yet wobble.

1. Mutagenic tautomerisation of the long A·G Watson-Crick DNA base mispair.

Tautomerisation processes of this long pair induce transfer of both A and G bases (but independently from each other) into the mutagenic A* and G* tautomeric forms, accordingly: at this four different H-bonded $A\cdot G^{*}_{\downarrow}(w)$, $A\cdot G^{*}_{\uparrow}(w)$, $A^{*}\cdot G_{\downarrow}(w)$ and $A^{*}\cdot G_{\uparrow}(w)$ pairs with wobble architecture are formed (Fig. 1). This number of the tautomerised pairs is not accidental – on the one hand, it is associated with the number of the acidic protons in each base pair, that are able to migrate one after another from one base to another (two of them represent itself amino proton at the N6 nitrogen atom of the A base and imino proton at the N1 nitrogen atom of the G base) forming in such a way a highly stable structure corresponding to the TS, and, on the other hand, with the number of the terminal, tautomerised wobble configurations involving mutagenic tautomers (this amount – 2 – can be also explained by the number of the DNA grooves, where one of the bases may shift relatively other base during this process of the mutagenic tautomerisation).

Among two different pathways of the mutagenic tautomerisation of the A·G(WC) DNA base mispair, in particular $A\cdot G(WC) \rightarrow A\cdot G^{*}_{\downarrow}(w)$ and $A\cdot G(WC) \rightarrow A\cdot G^{*}_{\uparrow}(w)$, first of them is

much faster than the other and therefore is appealing from the biological point of view (Figs. 1, 2 and Tables 1, 2, S1, S2). This A·G(WC)→A·G*$_{\downarrow}$(w) tautomeric conversion is initiated by the proton transfer at the N1 nitrogen atom of the G base along the intermolecular N1H···N1 H-bond to the N1 nitrogen atom of the A base. The $TS^{A^{+}\cdot G^{-}}{}_{A\cdot G(WC)\leftrightarrow A\cdot G^{*}\downarrow(w)}$ transition state, stabilized by the $N6^{+}H\cdots O6^{-}$ (7.46), $N1^{+}H\cdots O6^{-}$ (4.21) and $N1^{+}H\cdots N1^{-}$ (2.96 kcal·mol^{-1}) H-bonds in addition to the strong electrostatic interactions (Table S2), possesses wobble configuration due to the displacement of the deprotonated G^{-} base relatively the A^{+} base to the side of the DNA minor groove. Formed within this route A·G*$_{\downarrow}$(w) tautomerised base pair has shifted configuration and is stabilized by two N6H···O6 (1.64) and O6H···N1 (8.19 kcal·mol^{-1}) H-bonds. Interestingly, that herewith the O6H hydroxyl group of the G* base plays simultaneously the role of donor, as well as acceptor of the H-bonding (Table 1).

Among two others tautomerisation pathways (Tables 1, 2, S1, S2 and Figs. 1, 2), namely A·G(WC)→A*·G$_{\uparrow}$(w) and A·G(WC)→A*·G$_{\downarrow}$(w), second of them is extremely slow even in comparison with the time of the DNA replication in the cell (~10^{6} s [65,66]) and so it does not have actual biological meaning. The first A·G(WC)→A*·G$_{\uparrow}$(w) tautomerisation process starts with the migration of the proton localized at the N1 nitrogen atom of the G base along the intermolecular N1H···N1 H-bond to the N1 nitrogen atom of the A base. The $TS^{A^{+}\cdot G^{-}}{}_{A\cdot G(WC)\leftrightarrow A^{*}\cdot G\uparrow(w)}$ transition state, stabilized by the $N6^{+}H\cdots O6^{-}$ (2.53), $N6^{+}H\cdots N1^{-}$ (4.94), $N1^{+}H\cdots N1^{-}$ (2.29) and $N1^{+}H\cdots N2^{-}$ (4.88 kcal·mol^{-1}) H-bonds along with the strong electrostatic interactions (Table S2), as well as tautomerised A*·G$_{\uparrow}$(w) DNA base mispair, joined by the N1H···N2 (3.24), N1H···N6 (3.81) and N2H···N6 (3.95 kcal mol^{-1}) H-bonds, have wobble configurations due to the displacement of the deprotonated G^{-} base relatively the A^{+} base towards the DNA major groove (Table 1).

The A·G(WC)↔A·G*$_{\downarrow}$(w) and A·G(WC)↔A*·G$_{\uparrow}$(w) tautomerisation reactions are accompanied by the substantial rebuilding of the long A·G(WC) DNA base mispair from Watson-Crick to wobble sizes and *vice versa* that is displayed by the significant changes in its glycosidic parameters, namely the R(H_9-$H_{9'}$) glycosidic distances and the α_1 and α_2 glycosidic angles, varying largely without ruptures (Fig. S1). Both transition states of these reactions represent itself highly polar, highly stable $A^{+}\cdot G^{-}$ tight ion pairs shifted towards minor or major DNA grooves, respectively.

Also we have obtained 11 unique patterns of the specific intermolecular interactions including AH···B H-bonds and loosened A-H-B covalent bridges consistently replacing each other together with van der Waals contact during these tautomerisation processes (Fig. 3, Table 3). Two loosened A-H-B covalent bridges that are linked with H-bonds display the common feature of these transformations. Notably, that in the case of the A·G(WC)↔A·G*$_{\downarrow}$(w) pathway the N6H···O6 H-bond continuously exists along the whole IRC. The N2H···HC2 DH-bond [11,52] smoothly and without bifurcations transforms into the C2H···N2 H-bond in the course of the A·G(WC)↔A*·G$_{\uparrow}$(w) tautomerisation. Formation of the attractive N6···N2 van der Waals contact precedes the appearance of the N2H···N6 H-bond during the A·G(WC)↔A*·G$_{\uparrow}$(w) route. It has been established that the base pair that tautomerises remains in the A^{+}·G^{-} highly polar state, stabilized by rather strong electrostatic interactions, in the wide range of the IRC (-13.94÷4.17 and -13.21÷7.13 Bohr for the A·G(WC)↔A·G*$_{\downarrow}$(w) and A·G(WC)↔A*·G$_{\uparrow}$(w) conversions, respectively) (Fig. 3, Tables 3 and S2).

It is known that one of the necessary conditions for the successful incorporation of the incorrect DNA base pair into the structure of the DNA double helix is its flatness or (in the case of the non-planar structure) the ability to adapt its geometry to the planar [56,67,68].

To ensure that considered DNA base mispairs and transition states of their mutagenic tautomerisation with non-planar structures, involved in the biologically important processes of tautomerisation through the sequential DPT – namely, A·G(WC), A*·G$_{\uparrow}$(w), $TS^{A+\cdot G-}_{A\cdot G(WC)\leftrightarrow A^{*}\cdot G\uparrow(w)}$ and $TS^{A+\cdot G-}_{A\cdot G(WC)\leftrightarrow A\cdot G^{*}\downarrow(w)}$ – we investigated their structural non-rigidity, i.e. the ability of their mirror-symmetric conformers mutually transform (interconvert) into each other.

It has been found in all cases that the transition states of their interconversion are plane-symmetric structures, while in the case of the Watson-Crick-like A·G(WC) base mispair and $TS^{A+\cdot G-}_{A\cdot G(WC)\leftrightarrow A\cdot G^{*}\downarrow(w)}$ – plane complexes (Figs. S2, S3 and Table S3). This interconversion of the mirror-symmetric conformers of the slightly non-planar A·G(WC) base mispair and $TS^{A+\cdot G-}_{A\cdot G(WC)\leftrightarrow A\cdot G^{*}\downarrow(w)}$ transition state is realized by the mechanism of the planar inversion of the amino fragment of the G base. In all other cases the mechanism of the interconversion is reduced to the rotation of the amino group of the G base *via* the plane symmetric transition state, when its amine nitrogen atom N2 acts as a donor of the H-bonding (Figs. S2 and S3). It is characteristic that these processes occur without dissociation of the complexes into the monomers. For all

investigated structures their energy of planarization $\Delta\Delta G_{TS}$ is considerably less than the energy of the stacking interactions of the neighboring DNA base pairs [56], and these processes are quite fast ($\tau_{99.9\%}$=1.6·10^{-11}–1.2·10^{-7} s) (Table S3). This gives us a reason to believe that all studied structures have every chance to be integrated into the structure of the DNA double helix.

2. Mutagenic tautomerisation of the short C·T Watson-Crick DNA base mispair.

As in the case of the A·G(WC) mispair, the tautomerisation process of the C·T(WC) DNA base mispair into the wobble mismatches is also accompanied by the individual transitions of the C and T bases in its composition into the C* and T* mutagenic tautomeric forms, respectively. At this four different H-linked wobble C*·T$_{\downarrow}$(w), C*·T$_{\uparrow}$(w), C·T*$_{\downarrow}$(w) and C·T*$_{\uparrow}$(w) DNA base pairs are formed (Figs. 4 and 5). The C·T(WC)↔C*·T$_{\uparrow}$(w) and C·T(WC)↔C·T*$_{\downarrow}$(w) tautomerisation processes are much faster, than slow C·T(WC)↔C*·T$_{\downarrow}$(w) and C·T(WC)↔C·T*$_{\uparrow}$(w) transitions (Tables 4 and 5), that make no biological sense due to the reasons outlined above for the conversion of the A·G(WC) nucleobase mispair.

It is interesting to note, that the $TS^{C\text{-}\cdot T+}_{C\cdot T(WC)\leftrightarrow C\cdot T^{*}\uparrow(w)}$ transition state is stabilized except two O4^{+}H···N4^{-} (9.90) and N3^{+}H···N4^{-} (12.79) H-bonds, which are locked on the one and the same N4^{-} nitrogen atom, by the attractive N3^{+}···O2^{-} (0.48 kcal·mol^{-1}) van der Waals contact (Table 4).

Both of the biologically important C·T(WC)↔C*·T$_{\uparrow}$(w) and C·T(WC)↔C·T*$_{\downarrow}$(w) tautomerisation processes are at first initiated by the very fast ($\tau_{99.9\%}$=2.13·10^{-13} s) C·T(WC)→C*·T*(WC) double proton transfer along intermolecular H-bonds, that was previously investigated in details [12], followed by the formation of the C*·T*(WC) base pair, that in both cases is dynamically unstable intermediate (Table 5). Highly polar $TS^{C+\cdot T-}_{C^{*}\cdot T^{*}(WC)\leftrightarrow C^{*}\cdot T\uparrow(w)}$ and $TS^{C+\cdot T-}_{C^{*}\cdot T^{*}(WC)\leftrightarrow C\cdot T^{*}\downarrow(w)}$ transition states, that are stabilized in addition to the strong electrostatic interactions by the four N4^{+}H···O4^{-} (4.33), N4^{+}H···N3^{-} (2.93), N3^{+}H···N3^{-} (3.19), N3^{+}H···O2^{-} (3.25) and two N4^{+}H···O4^{-} (9.06), N3^{+}H···O4^{-} (7.52 kcal·mol^{-1}) H-bonds, respectively, is created by the displacement of the deprotonated T^{-} base relatively the protonated C^{+} base towards major and minor DNA grooves, correspondingly. Each of the terminal structures – tautomerised C*·T$_{\uparrow}$(w) (∠C4N3(C)C2N3(T)=0.0°) and C·T*$_{\downarrow}$(w)

(∠C4N3(C)C2N3(T)=34.3°) pairs with wobble geometry involving C* and T* mutagenic tautomers – are stabilized by two intermolecular N3H···N4 (7.45), N3H···O2 (5.13) and N4H···O4 (1.00), O4H···N3 (8.16 kcal·mol^{-1}) H-bonds, respectively.

Each of the $TS^{C+\cdot T-}_{C^*\cdot T^*(WC)\leftrightarrow C\cdot T^*\downarrow(w)}$ transition state and tautomerised C·T*$_{\downarrow}$(w) DNA base pair, involved in the biologically important process of the C·T(WC)↔C·T*$_{\downarrow}$(w) mutagenic tautomerisation, are united by two intermolecular H-bonds. In the first case, the O4^{-} oxygen atom of the T^{-} base is joint acceptor of the N4^{+}H···O4^{-} (9.06) and N3^{+}H···O4^{-} (7.52 kcal·mol^{-1}) H-bonds. In the latter case, the O4H hydroxyl group of the T* base acts simultaneously as donor and acceptor of the H-bonding – N4H···O4 (1.00) and O4H···N3 (8.16 kcal·mol^{-1}) (Table 1).

In fact, the DPT tautomerisation of the biologically important C*·T$_{\uparrow}$(w) nucleobase mispair into the C·T*$_{O2\uparrow}$(w) mismatch (Figs. S4, S5 and Table S4) does not occur, since the last mispair is dynamically unstable structure and any of its six intermolecular vibrations (21.7, 29.3, 59.8, 77.0, 95.3 and 102.9 cm^{-1}) can develop during its lifetime (τ=2.06·10^{-13} s). The planar C·T*$_{\downarrow}$(w) nucleobase mispair can be successfully incorporated into the structure of the DNA double helix, since energy of its planarization ($\Delta\Delta G_{TS}$=4.69 kcal·mol^{-1} under normal conditions (Table S3)) is noticeably smaller than the energy of the stacking interactions between Watson-Crick DNA base pairs [56].

Both of the C·T(WC)→C*·T$_{\uparrow}$(w) and C·T(WC)→C·T*$_{\downarrow}$(w) tautomerisation reactions, that are able to cause the C* and T* mutagenic tautomers, occur by the non-dissociative mechanism. This fact can be easily confirmed by the inseparable character of the profiles of the R(H_1-$H_{1'}$) glycosidic distances and α_1, α_2 glycosidic angles (Fig. S6).

Another distinguishing feature of these tautomeric conversions is their accompaniment by the unique patterns of the intermolecular interactions changing each other: 14 specific contacts including from 2 to 4 H-bonds and single attractive van der Waals contact for the C·T(WC)↔C*·T$_{\uparrow}$(w) process and 12 specific contacts, among which there are maximum 2 H-bonds and single attractive van der Waals contact, for the C·T(WC)↔C·T*$_{\downarrow}$(w) process. Common feature of these sweeps is the sequential interchange between the N4H···O4 and O4H···N4 H-bonds with their respective formation and disruption *via* the N4-H-O4 loosened covalent bridge (Fig. 6 and Table 6). Simultaneous existence of the N4H···O4/N3,

N3H···N3/O2 H-bonds within IRC=13.71÷42.02 Bohr and N4H···O4, N3H···O4 H-bonds within IRC=14.70÷42.02 Bohr for the C·T(WC)↔C*·T$_{\uparrow}$(w) and C·T(WC)↔C·T*$_{\downarrow}$(w) tautomerisation reactions, accordingly, reflects the existence of the highly polar $C^{+}\cdot T^{-}$ ion pair (Table 6).

3. Biological implication of the revealed routes of the mutagenic tautomerisation of the A·G and C·T DNA base mispairs with Watson-Crick shape.

Starting point of the study (now there are already all grounds to consider it as universal in the field of spontaneous point mutagenesis) is classical Watson-Crick tautomeric hypothesis, which now is already partially implemented at the level of "molecular phenomenology" [1,3,69,70].

It is characteristic that for all four biologically important, thermodynamically stable tautomerised A·G*$_{\downarrow}$(w), A*·G$_{\uparrow}$(w) and C*·T$_{\uparrow}$(w), C·T*$_{\downarrow}$(w) DNA base mispairs their lifetime τ is by many orders greater than the time (several nanoseconds [71]), necessary for the DNA-polymerase for the forced dissociation of the pairs during DNA replication. Moreover, the energy of the interaction between bases within these pairs is less (Table S2), than analogical value for the G·C Watson-Crick DNA base pair (29.28 kcal·mol^{-1} [14]). These two facts together mean that DNA-polymerase is able to successfully dissociate all these mispairs with the participation of the mutagenic tautomers into the monomers without changing their tautomeric status. In such a way, the A·G(WC)→A·G*$_{\downarrow}$(w), A·G(WC)→A*·G$_{\uparrow}$(w), C·T(WC)→C*·T$_{\uparrow}$(w) and C·T(WC)→C·T*$_{\downarrow}$(w) tautomerisation processes may be responsible for the generation of the mutagenic tautomers of the canonical DNA bases during DNA replication and are perfectly acceptable alternatives for the classical Löwdin's mechanism [19,20].

Moreover, results obtained in this paper allow us to understand at least at the semi-quantitative level, in what way the A·G/G·A and C·T/T·C non-equivalent [18] errors are formed during DNA replication within the framework of the classical tautomeric hypothesis. It turns out that mutagenic tautomer of each DNA base (A*, G*, C* and T*) belonging to the template can interact not only with the C, T, A and G DNA bases, accordingly, as it was previously thought [1,3,69,70], but also with the G, A, T and C bases, respectively, forming

aforementioned wobble A·G*$_{\downarrow}$(w), A*·G$_{\uparrow}$(w), C*·T$_{\uparrow}$(w) and C·T*$_{\downarrow}$(w) pairs with all the consequences arising therefrom.

All shifted thermodynamically stable (ΔG_{int}<0 (Table S2)) C*·T$_{\uparrow}$(w), C·T*$_{\downarrow}$(w) and A·G*$_{\downarrow}$(w), A*·G$_{\uparrow}$(w) DNA base mispairs may acquire enzymatically competent conformation tautomerising directly into the short Watson-Crick-like C·T(WC) base pair or through the A·G(WC)/G·A(WC) intermediates into the G·A$_{syn}$ and A*·G*$_{syn}$ pairs [72], respectively.

Conclusions.

It was uncovered for the first time novel mechanism of the mutagenic tautomerisation for the long A·G(WC) and short C·T(WC) Watson-Crick DNA base pairs. Tautomerisation of each mispair proceeds *via* four topologically and energetically different ways through highly stable transition states - H-bonded tight ion pairs comprising protonated and deprotonated bases.

Non-dissociative tautomerisation process is accompanied by the significant changing of the binding architecture from the Watson-Crick to wobble geometry both towards the minor, as well as major DNA grooves and *vice versa*. In all cases tautomerisation reactions are concerted and stepwise.

It was established for the first time that fast, biologically important processes of the mutagenic tautomerisation of the A·G(WC) and C·T(WC) base pairs are initiated by the transition of the more acidic, imino protons and is determined by the transition states of the $A^{+}\cdot G^{-}$ and $C^{+}\cdot T^{-}$ tight ion pair type. On the contrary, moving of the less acidic amino proton with the formation of the transition states of the $A^{-}\cdot G^{+}$ and $C^{-}\cdot T^{+}$ type causes too slow mutagenic tautomerisation that has no real biological meaning.

Biological significance of the obtained results is twofold. First, discovered for the first time phenomenon of the tautomerisation can be the source of the mutagenic tautomers of all four canonical nucleobases at the DNA replication, since wobble A·G*$_{\downarrow}$(w), A*·G$_{\uparrow}$(w), C*·T$_{\uparrow}$(w) and C·T*$_{\downarrow}$(w) tautomerised pairs meet all necessary for this requirements [71].

Second, these results allow us to understand the basic physico-chemical principles of the arising of some DNA replication errors caused by random mutagenic tautomerisation of the nucleotide bases belonging to the template. It becomes clear, in what way wobble C*·T(w), T*·C(w), A*·G(w) and G*·A(w) pairs (base belonging to the incoming nucleotide is placed on

the right) transform either into the short C·T(WC) and T·C(WC) Watson-Crick base pairs, or into the long A·G(WC) and G·A(WC) Watson-Crick base pairs. This knowledge is important, since high-fidelity DNA-polymerase selects base mispairs basing on their ability to acquire Watson–Crick-like configuration in its active site in the closed conformation state during the catalytic cycle [73-77]. The process of the acquisition of the enzymatically competent conformation in the hydrophobic recognition pocket of the high-fidelity DNA-polymerase has been earlier described by us in details [78].

Overall, elucidation of the novel mechanisms of the DNA bases tautomerisation eventually gives theoretical background for the experimental observations [73-77], allowing to suppose that DNA-polymerase is able to make errors due to the formation of the rare tautomers incorporated into structure of the wrong mispairs mimicking the shape of the canonical Watson-Crick pairs [79]. So, ultimately these mismatches can be accommodated without obstacles into the structure of the DNA double helix.

Acknowledgments. O.O.B. was supported by the grant of the President of Ukraine to support scientific research of young scientists for 2015 year from the State Fund for Fundamental Research of Ukraine (project № GP/F61/028) and by the grant of the National Academy of Sciences of Ukraine for scientific and research work of young scientists for 2015-2016 years. This work was performed using computational facilities of joint computer cluster of SSI "Institute for Single Crystals" of the National Academy of Sciences of Ukraine and Institute for Scintillation Materials of the National Academy of Sciences of Ukraine incorporated into Ukrainian National Grid. The authors sincerely thank Dr. Ivan S. Voiteshenko (Institute of High Technologies, Taras Shevchenko National University of Kyiv) and Dr. Fernando R. Clemente (Gaussian, Inc.) for their technical assistance of the work.

References.

1. Watson, J. D., & Crick, F. H. C. (1953). The Structure of DNA. *Cold Spring Harbor Symposia on Quantitative Biology, 18,* 123–31.
2. Freese, E.B. (1959). On the molecular explanation of spontaneous and induced mutations, *Brookhaven Symposium in Biology, 12*, 63-73.
3. Topal, M.D., & Fresco, J.R. (1976). Complementary base pairing and the origin of substitution mutations. *Nature, 263*, 285-289.
4. von Borstel, R.C. (1994). Origins of spontaneous base substitutions. *Mutation Research*, *307*, 131-140.
5. Brown, T., Hunter, W. H., Kneale, G., & Kennard, O. (1986). Molecular structure of the G·A base pair in DNA and its implications for the mechanism of transversion mutations. *Proceedings of the National Academy of Sciences of the United States of America*, *83*, 2402-2406.
6. Privé, G.G., Heinemann, U., Chandrasegaran, S., Kan, L.S., Kopka, M.L., & Dickerson, R.E. (1987). Helix geometry, hydration, and G·A mismatch in a B-DNA decamer. *Science*, *238*, 498-504.
7. Leonard, G.A., Booth, E.D., & Brown, T. (1990). Structural and thermodynamic studies on the adenine·guanine mismatch in B-DNA. *Nucleic Acids Research*, *18,* 5617-5623.
8. Boulard, Y., Cognet, J.A.H., & Fazakerley G.V. (1997). Solution structure as a function of pH of two central mismatches, C·T and C·C, in the 29 to 39 K-*ras* gene sequence, by nuclear magnetic resonance and molecular dynamics. *Journal of Molecular Biology*, *268*, 331-347.
9. Allawi, H.T., & SantaLucia, Jr. J. (1998). Thermodynamics of internal C·T mismatches in DNA. *Nucleic Acids Research*, *26*, 2694-2701.
10. Rossetti, G., Dans, P.D., Gomez-Pinto, I., Ivani, I., Gonzalez, G., & Orozco, M. (2015). The structural impact of DNA mismatches. *Nucleic Acids Research*, *43*, 4309-4321.
11. Brovarets' O.O., Zhurakivsky R.O., Hovorun D.M. (2014). Is the DPT tautomerisation of the long A·G Watson-Crick DNA base mispair a source of the adenine and guanine mutagenic tautomers? A QM and QTAIM response to the biologically important question. *Journal of Computational Chemistry, 35*, 451-466.
12. Brovarets', O.O., & Hovorun, D.M. (2013). Atomistic understanding of the C·T mismatched DNA base pair tautomerization *via* the DPT: QM and QTAIM computational approaches. *Journal of Computional Chemistry, 34*, 2577-2590.
13. Brovarets', O.O., & Hovorun, D.M. (2014). Can tautomerisation of the A·T Watson-Crick base pair *via* double proton transfer provoke point mutations during DNA replication? A comprehensive QM and QTAIM analysis. *Journal of Biomolecular Structure and Dynamics, 32*, 127-154.
14. Brovarets', O.O., & Hovorun, D.M. (2014). Why the tautomerization of the G·C Watson–Crick base pair *via* the DPT does not cause point mutations during DNA replication? QM and

QTAIM comprehensive analysis. *Journal of Biomolecular Structure & Dynamics, 32*, 1474-1499.

15. Kosenkov, D., Kholod, Y., Gorb, L., Shishkin, O., Hovorun, D.M., Mons, M., & Leszczynski, J. (2009). *Ab initio* kinetic simulation of gas-phase experiments: tautomerization of cytosine and guanine. *Journal of Physical Chemistry B*, *113*, 6140-6150.

16. Furmanchuk, A., Isayev, O., Gorb, L., Shishkin, O. V., Hovorun, D. M., & Leszczynski, J. (2011). Novel view on the mechanism of water-assisted proton transfer in the DNA bases: Bulk water hydration. *Physical Chemistry Chemical Physics, 13*, 4311‑4317.

17. Samijlenko, S.P., Yurenko, Y.P., Stepanyugin, A.V., & Hovorun, D.M. (2012). Tautomeric equilibrium of uracil and thymine in model protein−nucleic acid contacts. Spectroscopic and quantum chemical approach. *Journal of Physical Chemistry B, 114*, 1454-1461.

18. Huang, M.M., Arnheim, N, & Goodman, M.F. (1992). Extension of base mispairs by Taq DNA polymerase: implications for single nucleotide discrimination in PCR. *Nucleic Acids Research, 20,* 4567-4573.

19. Löwdin P.-O. Proton tunneling in DNA and its biological implications, *Rev. Mod. Phys.*, 1963, 35, 724-732.

20. Löwdin P.-O. Quantum genetics and the aperiodic solid: Some aspects on the biological problems of heredity, mutations, aging, and tumors in view of the quantum theory of the DNA molecule. In Löwdin, P.-O. (Ed.) *Advances in Quantum Chemistry* (1966, **2**, pp. 213-360). New York, USA, London, UK: Academic Press.

21. Brovarets', O. O., & Hovorun, D. M. (2015). Tautomeric transition between wobble A·C DNA base mispair and Watson-Crick-like A·C* mismatch: miscrostructural mechanism and biological significance. *Physical Chemistry Chemical Physics, 17*, 15103-15110.

22. Brovarets', O. O., & Hovorun, D. M. (2015). How many tautomerisation pathways connect Watson-Crick-like G*·T DNA base mispair and wobble mismatches? *Journal of Biomolecular Structure and Dynamics,* DOI: 10.1080/07391102.2015.1046936.

23. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., & Cheeseman, J.R., … Pople, J.A. (2010). *GAUSSIAN 09* (Revision B.01). Wallingford CT: Gaussian Inc.

24. Tirado-Rives, J., & Jorgensen, W.L. (2008). Performance of B3LYP Density Functional Methods for a large set of organic molecules. *Journal of Chemical Theory and Computation*, *4,* 297–306.

25. Parr, R.G., & Yang, W. (1989). *Density-functional theory of atoms and molecules*. Oxford: Oxford University Press.

26. Lee, C., Yang, W., & Parr, R.G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B: Condensed Matter and Materials Physics*, *37*, 785-789.

27. Matta, C.F. (2010). How dependent are molecular and atomic properties on the electronic structure method? Comparison of Hartree-Fock, DFT, and MP2 on a biologically relevant set of molecules. *Journal of Computational Chemistry, 31*, 1297–1311.
28. Brovarets', O. O., & Hovorun, D. M. (2015). The physicochemical essence of the purine·pyrimidine transition mismatches with Watson-Crick geometry in DNA: A·C* *versa* A*·C. A QM and QTAIM atomistic understanding. *Journal of Biomolecular Structure & Dynamics, 33,* 28-55.
29. Brovarets', O. O., & Hovorun, D. M. (2015). The nature of the transition mismatches with Watson-Crick architecture: the G*·T or G·T* DNA base mispair or both? A QM/QTAIM perspective for the biological problem. *Journal of Biomolecular Structure & Dynamics, 33,* 925-945.
30. Brovarets', O.O., Zhurakivsky, R.O., & Hovorun, D.M. (2013). The physico-chemical mechanism of the tautomerisation *via* the DPT of the long Hyp*·Hyp Watson-Crick base pair containing rare tautomer: a QM and QTAIM detailed look. *Chemical Physics Letters*, *578*, 126-132.
31. Brovarets', O.O., & Hovorun, D.M. (2014). How the long G•G* Watson-Crick DNA base mispair comprising keto and enol tautomers of the guanine tautomerises? The results of the QM/QTAIM investigation. *Physical Chemistry Chemical Physics, 6*, 15886-15899.
32. Peng, C., Ayala, P.Y., Schlegel, H.B., & Frisch, M.J. (1996). Using redundant internal coordinates to optimize equilibrium geometries and transition states. *Journal of Computational Chemistry, 17*, 49–56.
33. Frisch, M.J., Head-Gordon, M., & Pople, J.A. (1990). Semi-direct algorithms for the MP2 energy and gradient. *Chemical Physics Letters, 166,* 281-289.
34. Hariharan, P.C., & Pople, J.A. (1973). The influence of polarization functions on molecular orbital hydrogenation energies. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta), 28*, 213−222.
35. Krishnan, R., Binkley, J.S., Seeger, R., & Pople, J.A. (1980). Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *Journal of Chemical Physics, 72*, 650−654.
36. Kendall, R.A., Dunning, Jr., T.H., & Harrison, R.J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *Journal of Chemical Physics, 96*, 6796−6806.
37. Hratchian, H.P., & Schlegel, H.B. (2005). Finding minima, transition states, and following reaction pathways on ab initio potential energy surfaces. In Dykstra, C.E., Frenking, G., Kim, K.S., & Scuseria, G. (Eds.), *Theory and applications of computational chemistry: The first 40 years* (pp. 195-249). Amsterdam: Elsevier.
38. Brovarets', O.O., & Hovorun, D.M. (2010). Stability of mutagenic tautomers of uracil and its halogen derivatives: The results of quantum-mechanical investigation. *Biopolymers and Cell, 26*, 295–298.

39. Boys, S.F., & Bernardi, F. (1970). The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics*, *19,* 553–566.
40. Gutowski, M., Van Lenthe, J.H., Verbeek, J., Van Duijneveldt, F.B., & Chalasinski, G. (1986). The basis set superposition error in correlated electronic structure calculations. *Chemical Physics Letters, 124,* 370–375.
41. Sordo, J.A., Chin, S., & Sordo, T.L. (1988). On the counterpoise correction for the basis set superposition error in large systems. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta), 74*, 101–110.
42. Sordo, J.A. (2001). On the use of the Boys–Bernardi function counterpoise procedure to correct barrier heights for basis set superposition error. *Journal of Molecular Structure: THEOCHEM, 537,* 245–251.
43. Atkins, P.W. (1998). *Physical chemistry*. Oxford: Oxford University Press.
44. Wigner, E. (1932). Über das Überschreiten von Potentialschwellen bei chemischen Reaktionen [Crossing of potential thresholds in chemical reactions]. *Zeitschrift für Physikalische Chemie, B19*, 203−216.
45. Brovarets', O.O., & Hovorun, D.M. (2013). Atomistic nature of the DPT tautomerisation of the biologically important C•C* DNA base mispair containing amino and imino tautomers of the cytosine: A QM and QTAIM approach. *Physical Chemistry Chemical Physics, 15*, 20091-20104.
46. Brovarets', O.O., Zhurakivsky, R.O., & Hovorun, D.M. (2013). DPT tautomerization of the long A·A* Watson-Crick base pair formed by the amino and imino tautomers of adenine: combined QM and QTAIM investigation. *Journal of Molecular Modeling, 19*, 4223-4237.
47. Brovarets', O.O., Zhurakivsky, R.O., & Hovorun, D.M. (2014). Does the tautomeric status of the adenine bases change under the dissociation of the A*·A_{syn} Topal-Fresco DNA mismatch? A combined QM and QTAIM atomistic insight. *Physical Chemistry Chemical Physics, 16*, 3715-3725.
48. Bader, R.F.W. (1990). *Atoms in molecules: A quantum theory.* Oxford: Oxford University Press.
49. Keith, T.A. (2010). *AIMAll* (Version 10.07.01). Retrieved from aim.tkgristmill.com.
50. Brovarets', O. O., Yurenko, Y. P., & Hovorun, D. M. (2014). Intermolecular CH···O/N H-bonds in the biologically important pairs of natural nucleobases: A thorough quantum-chemical study. *Journal of Biomolecular Structure & Dynamics, 32,* 993-1022.
51. Brovarets', O. O., Yurenko, Y. P., & Hovorun, D. M. (2014). The significant role of the intermolecular CH···O/N hydrogen bonds in governing the biologically important pairs of the DNA and RNA modified bases: a comprehensive theoretical investigation. *Journal of Biomolecular Structure & Dynamics,* DOI: 10.1080/07391102.2014.968623.
52. Matta, C.F., Hernández-Trujillo, J., Tang, T.H., Bader, R.F. (2003). Hydrogen-hydrogen bonding: a stabilizing interaction in molecules and crystals. *Chemistry, 9*, 1940-1951.

53. Espinosa, E., Molins, E., & Lecomte, C. (1998). Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chemical Physics Letters*, *285*, 170–173.
54. Mata, I., Alkorta, I., Espinosa, E., & Molins, E. (2011). Relationships between interaction energy, intermolecular distance and electron density properties in hydrogen bonded complexes under external electric fields. *Chemical Physics Letters, 507*, 185–189.
55. Nikolaienko, T.Y., Bulavin, L.A., & Hovorun, D.M. (2012). Bridging QTAIM with vibrational spectroscopy: The energy of intramolecular hydrogen bonds in DNA-related biomolecules. *Physical Chemistry Chemical Physics, 14*, 7441–7447.
56. Iogansen, A.V. (1999). Direct proportionality of the hydrogen bonding energy and the intensification of the stretching *v*(XH) vibration in infrared spectra. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 55,* 1585–1612.
57. Saenger, W. (1984). *Principles of nucleic acid structure.* New York: Springer.
58. Russo, N, Toscano, M., & Grand, A. (2001). Bond energies and attachments sites of sodium and potassium cations to DNA and RNA nucleic acid bases in the gas phase. *Journal of the American Chemical Society*, *123*, 10272-10279.
59. Yang, Z., & Rodgers, M.T. (2004). Theoretical studies of the unimolecular and bimolecular tautomerization of cytosine. *Physical Chemistry Chemical Physics, 6*, 2749-2757.
60. Herrera, B., & Toro-Labbe, A. (2004). The role of the reaction force to characterize local specific interactions that activate the intramolecular proton transfers in DNA bases. *Journal of Chemical Physics, 121*, 7096- 7102.
61. Gorb, L., Kaczmarek, A., Gorb, A., Sadlej, A.J., & Leszczynski, J. (2005). Thermodynamics and kinetics of intramolecular proton transfer in guanine. Post Hartree–Fock study. *Journal of Physical Chemistry B*, *109*, 13770–13776.
62. Zhen-Min, Z., Qi-Ren, Z., & Chun-Yuan, G. (2006). Potential energy surface of cytosine and tunneling between its normal and trans-imino tautomer. *Communications in Theoretical Physics, 46*, 541-544.
63. Zhao, Z.-M., Zhang, Q.-R., Gao, C.-Y. & Zhuo, Y.-Z. (2006). Motion of the hydrogen bond proton in cytosine and the transition between its normal and imino states. *Physics Letters A, 359*, 10–13.
64. Brovarets', O.O., & Hovorun, D.M. (2010). How stable are the mutagenic tautomers of DNA bases? *Biopolymers and Cell, 26,* 72–76.
65. Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A., & Ellenberger, T. (2006). *DNA repair and mutagenesis*. Washington D.C.: ASM Press.
66. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. Chapter 5. The Initiation and Completion of DNA Replication in Chromosomes, *Molecular Biology of the Cell* (4th edition), ISBN-10: 0-8153-3218-1I, SBN-10: 0-8153-4072-9, New York: Garland Science; 2002.

67. Govorun, D.N., Danchuk, V.D., Mishchuk, Ya.R., Kondratyuk, I.V., Radomsky, N.F., & Zheltovsky, N.V. (1992). AM1 calculation of the nucleic acid bases structure and vibrational spectra. *Journal of Molecular Structure*, *267*, 99–103.
68. Nikolaienko, T.Y., Bulavin, L.A., & Hovorun, D.M. (2011). How flexible are DNA constituents? The quantum-mechanical study. *Journal of Biomolecular Structure & Dynamics*, *29*, 563-575.
69. Danilov, V.I., Anisimov, V.M., Kurita, N., & Hovorun, D. (2005). MP2 and DFT studies of the DNA rare base pairs: the molecular mechanism of the spontaneous substitution mutations conditioned by tautomerism of bases. *Chemical Physics Letters, 412*, 285-293.
70. Fonseca Guerra, C., Bickelhaupt, F.M., Saha, S., & Wang, F. (2006). Adenine tautomers: relative stabilities, ionization energies, and mismatch with cytosine. *Journal of Physical Chemistry A*, *110*, 4012-4020.
71. Brovarets', O. O., Kolomiets', I. M., & Hovorun, D. M. (2012). Elementary molecular mechanisms of the spontaneous point mutations in DNA: A novel quantum-chemical insight into the classical understanding. In T. Tada (Ed.), *Quantum chemistry – molecules for innovations* (pp. 59–102). Rijeka: In Tech Open Access.
72. Brovarets', O.O., & Hovorun, D.M. (2014). DPT tautomerisation of the G·A_{syn} and A*·G^*_{syn} DNA mismatches: A QM/QTAIM combined atomistic investigation. *Physical Chemistry Chemical Physics*, *16*, 9074-9085.
73. Bebenek, K., Pedersen, L. C., & Kunkel, T. A. (2011). Replication infidelity *via* a mismatch with Watson−Crick geometry. *Proceeding of the National Academy of Sciences of the United States of America, 108*, 1862−1867.
74. Christian, T. D., Romano, L. J., & Rueda, D. (2009). Single-molecule measurements of synthesis by DNA polymerase with base-pair resolution. *Proceedings of the National Academy of Sciences*, 106, 21109–21114.
75. Alemán, E.A., de Silva, C., Patrick, E.M., Musier-Forsyth, K., & Rueda, D. (2014). Single-molecule fluorescence using nucleotide analogs: A proof-of-principle. *Journal of Physical Chemistry Letters, 5*, 777-781.
76. Xia, S., Wang, J., & Konigsberg, W. H. (2013). DNA mismatch synthesis complexes provide insights into base selectivity of a B family DNA polymerase. *Journal of the American Chemical Society, 135*, 193−202.
77. Wang, W., Hellinga, H. W., and Beese, L. S. (2011). Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proceeding of the National Academy of Sciences of the United States of America, 108*, 17644−17648.
78. Brovarets', O.O., & Hovorun, D.M. (2014). Does the G•G^*_{syn} DNA mismatch containing canonical and rare tautomers of the guanine tautomerise through the DPT? A QM/QTAIM microstructural study. *Molecular Physics, 112*, 3033-3046.
79. Tsai, M.-D. (2014). How DNA polymerases catalyze DNA replication, repair, and mutation. *Biochemistry, 53*, 2749-2751.

Table 1. Electron-topological, structural, vibrational and energetic characteristics of the intermolecular H-bonds and DH-bonds in the DNA base mispairs containing A and G nucleobases and TSs of their tautomerisation *via* the sequential DPT and structural displacement of the A and G bases relative each other, energetic and polar characteristics of the latters obtained at the B3LYP/6-311++G(d,p) level of theory.

Base pair/TS	AH···B H-bond/DH-bond	ρ[a]	$\Delta\rho$[b]	$100\cdot\varepsilon$[c]	$d_{A\cdots B}$[d]	$d_{H\cdots B}$[e]	Δd_{AH}[f]	∠AH···B[g]	$\Delta\nu$[h]	$E_{AH\cdots B}$[i]	ΔG[j]	μ[k]
A·G(WC)[£]	N6H···O6	0.032	0.109	3.79	2.866	1.842	0.019	176.6	336.2	5.68	0.00	5.21
	N1H···N1	0.032	0.084	6.64	2.972	1.936	0.024	178.9	429.1	6.51		
	N2H···HC2	0.004	0.014	33.40	3.153	2.469	0.00007	124.6	-0.5	0.68*		
A·G*$_{\downarrow}$(w)	N6H···O6	0.012	0.048	5.41	3.043	2.246	0.004	134.6	64.6	1.64	3.76	2.39
	O6H···N1	0.047	0.103	4.89	2.695	1.755	0.034	154.5	655.8	8.19		
A·G*$_{\uparrow}$(w)	N6H···N1	0.023	0.070	6.74	3.083	2.065	0.014	176.0	255.8	4.85	6.28	2.87
	N2H···N1	0.024	0.073	7.10	3.061	2.046	0.014	172.6	258.0	4.87		
A*·G$_{\uparrow}$(w)	N1H···N6	0.020	0.059	6.20	3.087	2.157	0.010	150.3	173.1	3.81	14.29	8.38
	N2H···N6	0.021	0.065	6.58	2.991	2.154	0.011	138.0	183.0	3.95		
	N1H···N2	0.017	0.051	10.26	3.136	2.230	0.008	147.2	136.7	3.24		
$TS^{A+\cdot G-}_{A\cdot G(WC)\leftrightarrow A\cdot G^*\downarrow(w)}$	$N6^+H\cdots O6^-$	0.048	0.162	5.72	2.605	1.648	0.034	150.2	551.2	7.46	17.01	9.26
	$N1^+H\cdots O6^-$	0.029	0.091	10.22	2.800	1.946	0.033	136.4	512.3	4.21		
	$N1^+H\cdots N1^-$	0.023	0.072	18.76	3.011	2.045	0.033	152.1	512.3	2.96		
A*·G$_{\downarrow}$(w)	N1H···O6	0.015	0.057	0.81	2.973	2.185	0.005	132.9	71.5	1.85	21.11	9.54
	C2H···O6	0.009	0.034	25.94	3.146	2.488	-0.002	118.0	-29.4	1.75*		
$TS^{A+\cdot G-}_{A\cdot G(WC)\leftrightarrow A^*\cdot G\uparrow(w)}$	$N6^+H\cdots O6^-$	0.018	0.064	72.04	2.945	2.150	0.031	131.6	551.2	2.53	25.29	12.56
	$N6^+H\cdots N1^-$	0.033	0.091	2.01	2.866	1.904	0.031	152.3	551.2	4.94		
	$N1^+H\cdots N1^-$	0.014	0.049	99.24	3.131	2.324	0.023	133.7	512.3	2.29		
	$N1^+H\cdots N2^-$	0.028	0.076	2.49	2.984	1.981	0.023	150.2	512.3	4.88		
$TS^{A-\cdot G+}_{A\cdot G(WC)\leftrightarrow A\cdot G^*\uparrow(w)}$	$O6^+H\cdots N6^-$	0.050	0.090	7.51	2.719	1.740	0.055	158.9	1005.1	10.25	25.79	5.29
	$N1^+H\cdots N6^-$	0.068	0.093	7.94	2.642	1.617	0.071	155.5	1067.3	10.58		
	$N2^+H\cdots N1^-$	0.024	0.065	7.48	3.063	2.089	0.021	157.3	345.8	5.77		
$TS^{A-\cdot G+}_{A\cdot G(WC)\leftrightarrow A^*\cdot G\downarrow(w)}$	$O6^+H\cdots N1^-$	0.056	0.079	2.79	2.715	1.686	0.078	167.3	1330.8	11.86	49.20	12.87
	$N1^+H\cdots N1^-$	0.021	0.071	16.40	2.988	2.106	0.008	143.2	123.3	3.01		
	$N2^+H\cdots HC2^-$	0.005	0.016	36.78	3.083	2.277	0.0004	136.1	10.0	0.85*		

[a]The electron density at the (3,-1) BCP of the H-bond, a.u.

[b]The Laplacian of the electron density at the (3,-1) BCP of the H-bond, a.u.

[c]The ellipticity at the (3,-1) BCP of the H-bond

[d]The distance between the A (H-bond donor) and B (H-bond acceptor) atoms of the AH···B H-bond, Å

[e]The distance between the H and B atoms of the AH···B H-bond, Å

[f]The elongation of the H-bond donating group AH upon the AH···B H-bonding, Å

[g]The H-bond angle, degree

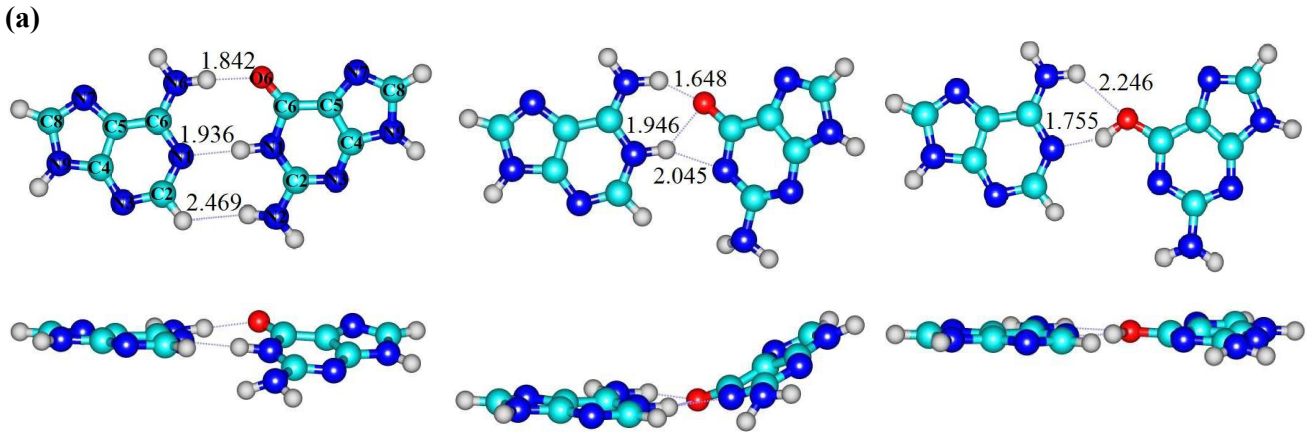
[h]The redshift of the stretching vibrational mode ν(AH) of the AH H-bonded group, cm^{-1}

[i]Energy of the H-bonds, calculated by Iogansen's [56] or Espinose-Molins-Lecomte (marked with an asterisk) [53,54] formulas, $kcal\cdot mol^{-1}$

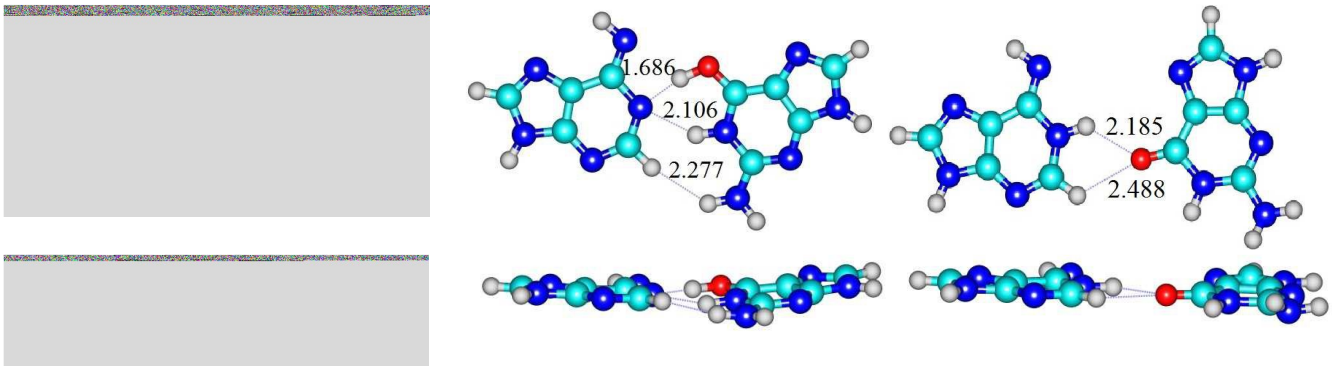
[j]The relative Gibbs free energy of the complex obtained at the MP2/cc-pVQZ//B3LYP/6-311++G(d,p) level of theory under normal conditions, $kcal\cdot mol^{-1}$

[k]The dipole moment of the complex, D

[£]Data are taken from the work [11].

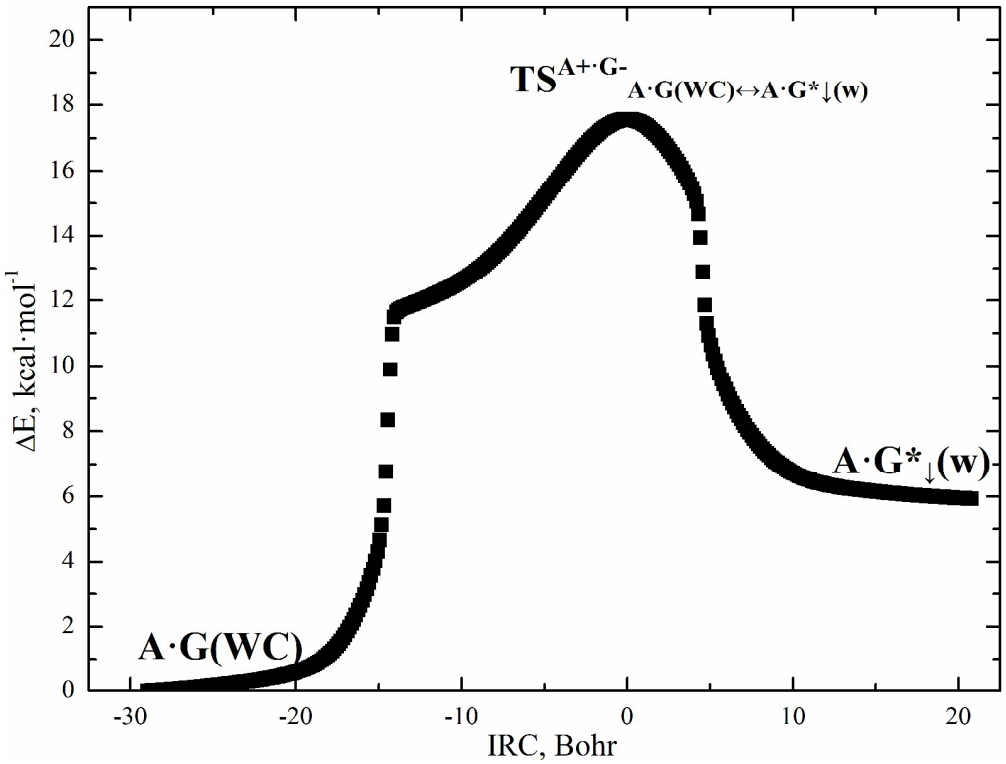


(d)

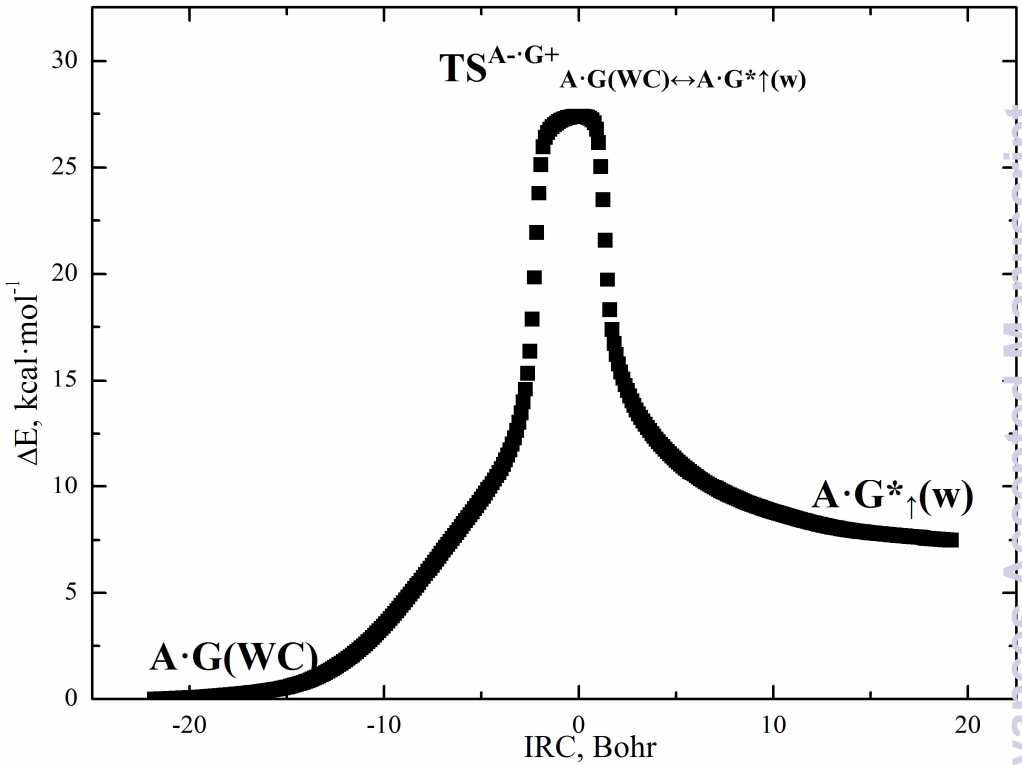


A·G(WC) **$TS^{A-\cdot G+}_{A\cdot G(WC)\leftrightarrow A^{*}\cdot G_{\downarrow}(w)}$** ($\nu_i$=-193.5*i* cm^{-1}) **$A^{*}\cdot G_{\downarrow}(w)$**

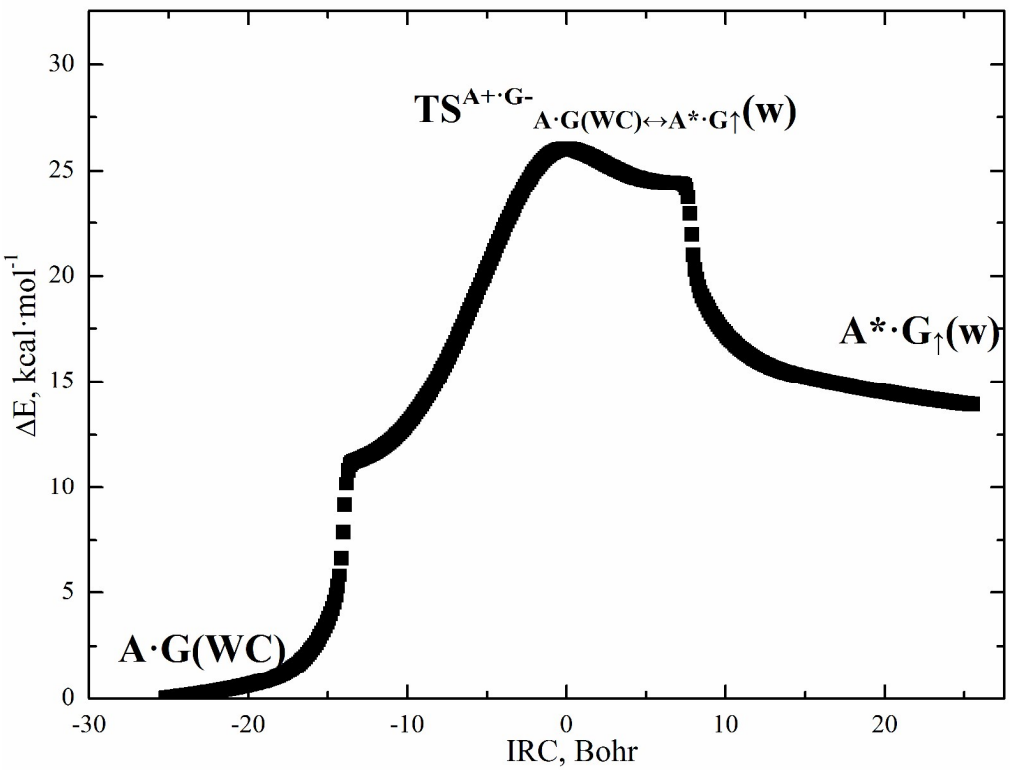
Fig. 1. Structures corresponding to the stationary points on the reaction pathways of the (**a**) A·G(WC)↔A·G*$_{\downarrow}$(w), (**b**) A·G(WC)↔A·G*$_{\uparrow}$(w), (**c**) A·G(WC)↔A*·G$_{\uparrow}$(w) and (**d**) A·G(WC)↔A*·G$_{\downarrow}$(w) conversions *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory. Dotted lines indicate AH···B H-bonds and AH···HB DH-bonds (their lengths are presented in angstroms). Carbon atoms are in light-blue, nitrogen – in dark-blue, hydrogen – in grey and oxygen – in red. ν_i – imaginary frequency.



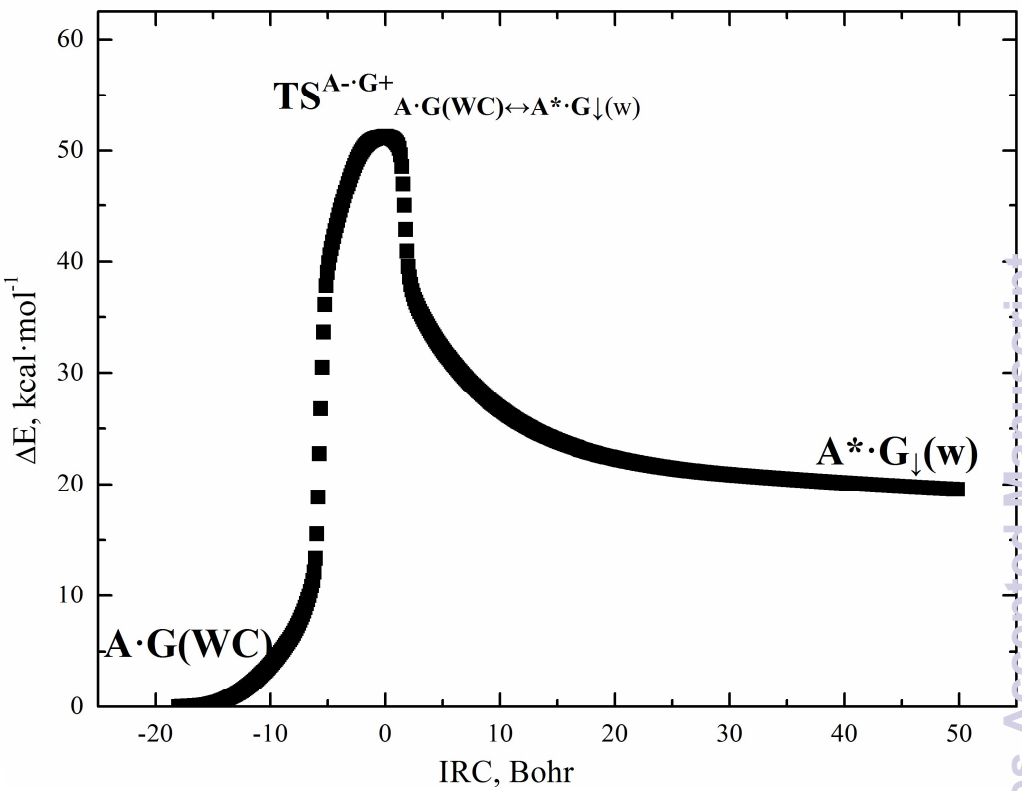
(a)



(b)



(c)



(d)

Fig. 2. Profiles of the relative electronic energy ΔE of the DNA base mispairs along the IRC of the (**a**) A·G(WC)↔A·G*↓(w), (**b**) A·G(WC)↔A·G*↑(w), (**c**) A·G(WC)↔A*·G↑(w) and (**d**) A·G(WC)↔A*·G↓(w) tautomerisations *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory.

Table 2. Energetic and kinetic characteristics of the A·G(WC)↔A·G*$_{\downarrow}$(w), A·G(WC)↔A·G*$_{\uparrow}$(w), A·G(WC)↔A*·G$_{\uparrow}$(w) and A·G(WC)↔A*·G$_{\downarrow}$(w) tautomerisations *via* the sequential DPT obtained at the different levels of theory for the geometry calculated at the B3LYP/6-311++G(d,p) level of theory (see also Fig. 1 and Table S1).

Level of theory	ΔG^a	ΔE^b	$\Delta\Delta G_{TS}^c$	$\Delta\Delta E_{TS}^d$	$\Delta\Delta G^e$	$\Delta\Delta E^f$	$\tau_{99.9\%}^g$
A·G(WC)↔A·G*$_{\downarrow}$(w)							
MP2/6-311++G(2df,pd)	3.27	5.70	16.88	16.94	13.60	11.24	$9.71\cdot10^{-3}$
MP2/6-311++G(3df,2pd)	3.63	6.06	16.79	16.86	13.16	10.79	$4.60\cdot10^{-3}$
MP2/cc-pVTZ	3.88	6.31	17.73	17.79	13.84	11.48	$1.45\cdot10^{-2}$
MP2/cc-pVQZ	3.76	6.19	17.01	17.07	13.25	10.88	$5.31\cdot10^{-3}$
A·G(WC)↔A·G*$_{\uparrow}$(w)							
MP2/6-311++G(2df,pd)	6.14	7.34	25.48	25.37	19.34	18.03	$1.52\cdot10^{2}$
MP2/6-311++G(3df,2pd)	6.36	7.56	25.97	25.86	19.61	18.29	$2.37\cdot10^{2}$
MP2/cc-pVTZ	6.34	7.54	25.35	25.23	19.01	17.69	$1.15\cdot10^{2}$
MP2/cc-pVQZ	6.28	7.48	25.79	25.68	19.51	18.20	$2.02\cdot10^{2}$
A·G(WC)↔A*·G$_{\uparrow}$(w)							
MP2/6-311++G(2df,pd)	14.16	13.96	24.95	24.05	10.79	10.09	$8.45\cdot10^{-5}$
MP2/6-311++G(3df,2pd)	14.11	13.91	25.01	24.11	10.90	10.20	$1.02\cdot10^{-4}$
MP2/cc-pVTZ	14.03	13.83	25.46	24.56	11.43	10.73	$2.49\cdot10^{-4}$
MP2/cc-pVQZ	14.29	14.09	25.29	24.39	11.00	10.30	$1.21\cdot10^{-4}$
A·G(WC)↔A*·G$_{\downarrow}$(w)							
MP2/6-311++G(2df,pd)	21.01	22.54	49.23	49.70	28.22	27.16	$4.66\cdot10^{8}$
MP2/6-311++G(3df,2pd)	20.73	22.26	48.94	49.41	28.21	27.15	$4.57\cdot10^{8}$
MP2/cc-pVTZ	21.39	22.92	49.53	50.00	28.14	27.08	$4.05\cdot10^{8}$
MP2/cc-pVQZ	21.11	22.6	49.20	49.67	28.08	27.02	$3.67\cdot10^{8}$

[a]The Gibbs free energy of the product relatively the reactant of the tautomerisation reaction (T=298.15 K), kcal·mol^{-1}

[b]The electronic energy of the product relatively the reactant of the tautomerisation reaction, kcal·mol^{-1}

[c]The Gibbs free energy barrier for the forward reaction of tautomerisation, kcal·mol^{-1}

[d]The electronic energy barrier for the forward reaction of tautomerisation, kcal·mol^{-1}

[e]The Gibbs free energy barrier for the reverse reaction of tautomerisation, kcal·mol^{-1}

[f]The electronic energy barrier for the reverse reaction of tautomerisation, kcal·mol^{-1}

[g]The time necessary to reach 99.9%of the equilibrium concentration between the reactant and the product of the tautomerisation reaction, s

See also summary Table S1 for the Gibbs and electronic energies of the mispairs and TSs relatively the global minimum – the long Watson-Crick-like A·G(WC) DNA base mispair.

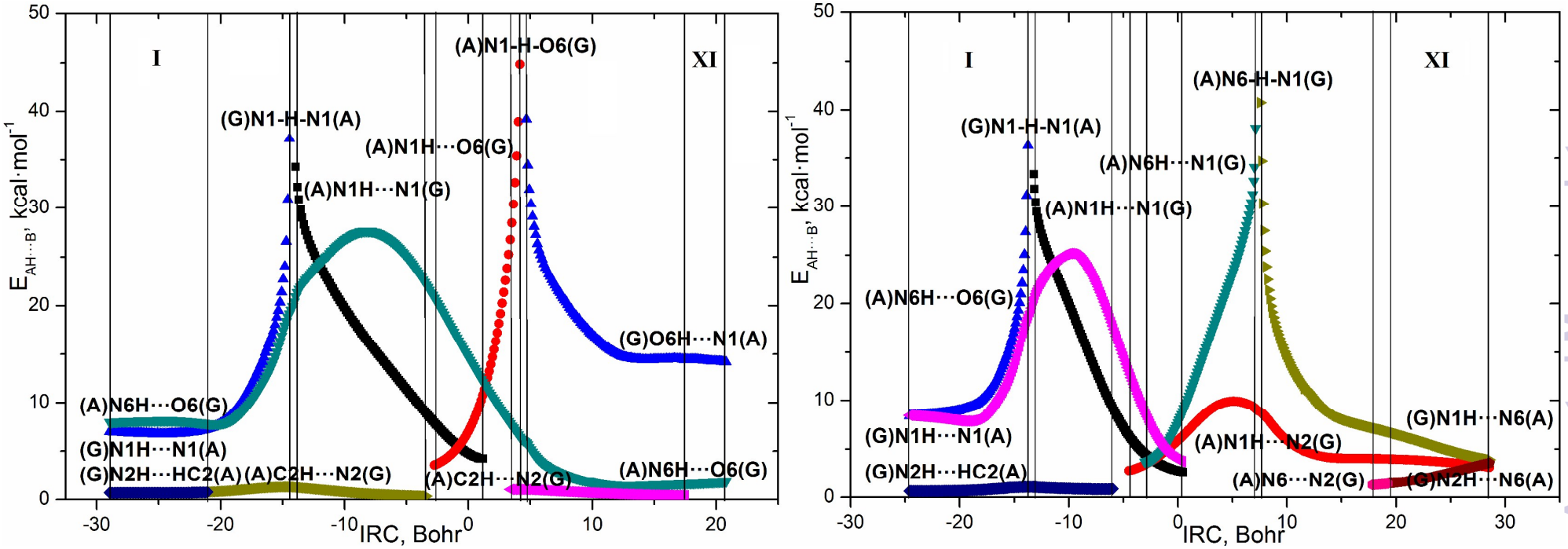


Fig. 3. Exchange of the patterns of the intermolecular AH···B H-bonds (their energies $E_{AH\cdots B}$ are estimated by the EML formula at the (3,-1) BCPs) along the IRC of the (**a**) A·G(WC)↔A·G*$_{\downarrow}$(w) and (**b**) A·G(WC)↔A*·G$_{\uparrow}$(w) biologically important tautomerisations *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory (see Tables 1 and 3).

Table 3. Patterns of the intermolecular interactions including AH···B H-bonds and loosened A-H-B covalent bridges that sequentially replace each other along the IRC of the A·G(WC)↔A·G*$_{\downarrow}$(w) and A·G(WC)↔A*·G$_{\uparrow}$(w) biologically important tautomerisations *via* the sequential DPT and their ranges of the existence obtained at the B3LYP/6-311++G(d,p) level of theory (see Figs. 1-3).

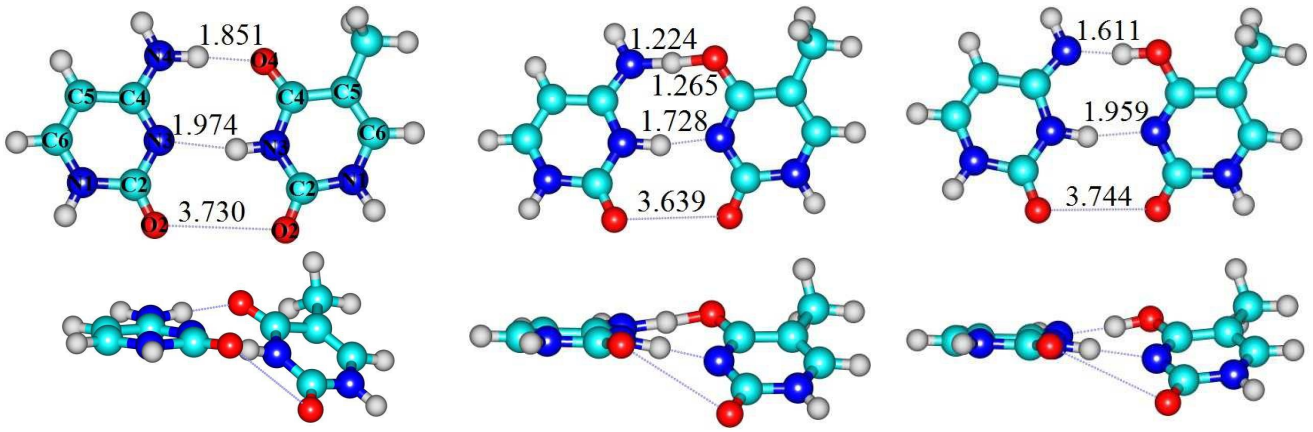
Patterns	IRC range, Bohr	Intermolecular interactions, forming patterns
A·G(WC)↔A·G*$_{\downarrow}$(w)		
I	[-28.96÷-21.06)	(A)N6H···O6(G), (G)N1H···N1(A), (G)N2H···HC2(A)
II	[-21.06÷-14.44)	(A)N6H···O6(G), (G)N1H···N1(A), (A)C2H···N2(G)
III	[-14.44÷-13.94)	(G)N1-H-N1(A), (A)N6H···O6(G), (A)C2H···N2(G)
IV	[-13.94÷-3.39)	(A)N6H···O6(G), (A)N1H···N1(G), (A)C2H···N2(G)
V	[-3.39÷-2.74)	(A)N6H···O6(G), (A)N1H···N1(G)
VI	[-2.74÷1.17)	(A)N6H···O6(G), (A)N1H···N1(G), (A)N1H···O6(G)
VII	[1.17÷3.39)	(A)N6H···O6(G), (A)N1H···O6(G)
VIII	[3.39÷4.17)	(A)N6H···O6(G), (A)N1H···O6(G), (A)C2H···N2(G)
IX	[4.17÷4.69)	(A)N1-H-O6(G), (A)N6H···O6(G), (A)C2H···N2(G)
X	[4.69÷17.54)	(G)O6H···N1(A), (A)N6H···O6(G), (A)C2H···N2(G)
XI	[17.54÷20.80]	(G)O6H···N1(A), (A)N6H···O6(G)
A·G(WC)↔A*·G$_{\uparrow}$(w)		
I	[-24.61÷-13.75)	(A)N6H···O6(G), (G)N1H···N1(A), (G)N2H···HC2(A)
II	[-13.75÷-13.21)	(G)N1-H-N1(A), (A)N6H···O6(G), (G)N2H···HC2(A)
III	[-13.21÷-5.96)	(A)N6H···O6(G), (A)N1H···N1(G), (G)N2H···HC2(A)
IV	[-5.96÷-4.50)	(A)N6H···O6(G), (A)N1H···N1(G)
V	[-4.50÷-3.04)	(A)N6H···O6(G), (A)N1H···N1(G), (A)N1H···N2(G)
VI	[-3.04÷0.45)	(A)N6H···O6(G), (A)N6H···N1(G), (A)N1H···N1(G), (A)N1H···N2(G)
VII	[0.45÷7.13)	(A)N6H···N1(G), (A)N1H···N2(G)
VIII	[7.13÷7.55)	(A)N6-H-N1(G), (A)N1H···N2(G)
IX	[7.55÷17.80)	(G)N1H···N6(A), (A)N1H···N2(G)
X	[17.80÷19.25)	(G)N1H···N6(A), (A)N1H···N2(G), (A)N6···N2(G)
XI	[19.25÷28.57]	(G)N1H···N6(A), (A)N1H···N2(G), (G)N2H···N6(A)

Table 4. Electron-topological, structural, vibrational and energetic characteristics of the intermolecular H-bonds and vdW contacts revealed in the DNA base mispairs containing C and T nucleobases and TSs of their mutual transformations *via* the sequential DPT and structural displacement of the C and T bases relative each other, energetic and polar characteristics of the latters obtained at the B3LYP/6-311++G(d,p) level of theory.

Base pair/TS	AH···B H-bond/A···B vdW contact	ρ[a]	$\Delta\rho$[b]	$100\cdot\varepsilon$[c]	$d_{A\cdots B}$[d]	$d_{H\cdots B}$[e]	Δd_{AH}[f]	∠AH···B[g]	$\Delta\nu$[h]	$E_{AH\cdots B/O\cdots O/N}$[i]	ΔG[j]	μ[k]
C·T(WC)[£]	N4H···O4	0.031	0.108	3.48	2.873	1.851	0.016	174.8	287.3	5.19	0.00	4.05
	N3H···N3	0.028	0.078	6.50	3.001	1.974	0.024	170.1	408.2	6.33		
	O2···O2	0.002	0.008	21.46	3.730	-	-	-	-	0.32*		
C*·T$_{\uparrow}$(w)	N3H···N4	0.039	0.095	6.30	2.877	1.838	0.032	172.4	549.4	7.45	0.52	1.08
	N3H···O2	0.028	0.100	4.31	2.906	1.882	0.017	173.0	282.1	5.13		
C*·T$_{\downarrow}$(w)	N3H···O4	0.027	0.097	2.86	2.912	1.902	0.016	166.8	255.1	4.84	3.08	0.54
	N3H···O2	0.029	0.106	3.75	2.878	1.860	0.017	169.6	277.0	5.08		
C·T*$_{O2}$(w)	N4H···N3	0.038	0.097	7.66	2.874	1.849	0.027	170.5	485.6	6.97	6.25	2.06
	O2H···N3	0.059	0.093	5.52	2.684	1.659	0.058	175.8	1088.0	10.68		
TS$_{C^*\cdot T\uparrow(w)\leftrightarrow C\cdot T^*O2(w)}$	N4H···N3	0.060	0.097	6.07	2.728	1.666	-	172.6	-	11.49**	6.67	5.08
C*·T*(WC)[£]	O4H···N4	0.066	0.096	4.69	2.637	1.611	0.060	173.9	1129.2	10.89	9.13	3.98
	N3H···N3	0.029	0.081	6.32	2.974	1.959	0.024	165.9	409.7	6.34		
	O2···O2	0.002	0.008	26.46	3.744	-	-	-	-	0.30*		
TS$_{C\cdot T(WC)\leftrightarrow C^*\cdot T^*(WC)}$[£]	N3H···N3	0.052	0.092	5.95	2.781	1.728	-	167.6	-	9.67**	9.53	5.86
	O2···O2	0.002	0.009	17.67	3.639	-	-	-	-	0.41*		
C·T*$_{\downarrow}$(w)	N4H···O4	0.012	0.042	23.47	3.120	2.312	0.003	136.2	49.2	1.00	12.05	7.36
	O4H···N3	0.045	0.099	3.29	2.727	1.774	0.034	157.2	651.6	8.16		
C·T*$_{\uparrow}$(w)	O4H···N4	0.026	0.074	6.84	2.966	2.002	0.013	166.3	266.8	4.97	15.69	7.10
	N4H···N3	0.015	0.050	17.23	3.032	2.316	0.010	126.4	143.1	3.35		
TS$^{C+\cdot T-}_{C^*\cdot T^*(WC)\leftrightarrow C^*\cdot T\uparrow(w)}$	N4$^+$H···O4$^-$	0.032	0.097	5.23	2.821	1.866	0.033	150.5	523.7	4.33	17.03	8.32
	N4$^+$H···N3$^-$	0.024	0.081	30.73	2.913	2.040	0.033	139.4	523.7	2.93		
	N3$^+$H···N3$^-$	0.020	0.064	15.00	3.014	2.138	0.027	140.1	421.0	3.19		
	N3$^+$H···O2$^-$	0.020	0.061	14.61	3.054	2.078	0.027	154.8	421.0	3.25		
TS$^{C+\cdot T-}_{C^*\cdot T^*(WC)\leftrightarrow C\cdot T^*\downarrow(w)}$	N4$^+$H···O4$^-$	0.062	0.174	5.42	2.538	1.558	0.049	151.2	793.2	9.06	26.62	12.16
	N3$^+$H···O4$^-$	0.036	0.105	5.12	2.731	1.840	0.036	140.1	559.9	7.52		
TS$^{C-\cdot T+}_{C\cdot T(WC)\leftrightarrow C^*\cdot T\downarrow(w)}$	O4$^+$H···N4$^-$	0.037	0.082	3.70	2.765	1.837	0.059	148.3	957.5	5.82	34.39	5.92
	O4$^+$H···N3$^-$	0.027	0.079	27.99	2.848	1.982	0.059	140.1	957.5	4.18		
	N3$^+$H···N3$^-$	0.018	0.067	131.66	2.977	2.175	0.039	131.1	618.5	2.55		
	N3$^+$H···O2$^-$	0.035	0.094	0.92	2.854	1.835	0.039	161.1	618.5	5.38		
TS$^{C-\cdot T+}_{C\cdot T(WC)\leftrightarrow C\cdot T^*\uparrow(w)}$	O4$^+$H···N4$^-$	0.054	0.107	6.50	2.615	1.678	0.051	150.7	940.9	9.90	46.74	10.54
	N3$^+$H···N4$^-$	0.086	0.046	5.77	2.617	1.531	0.136	154.3	1541.6	12.79		
	O2$^+$···N3$^-$	0.003	0.010	118.00	3.684	-	-	-	-	0.48*		

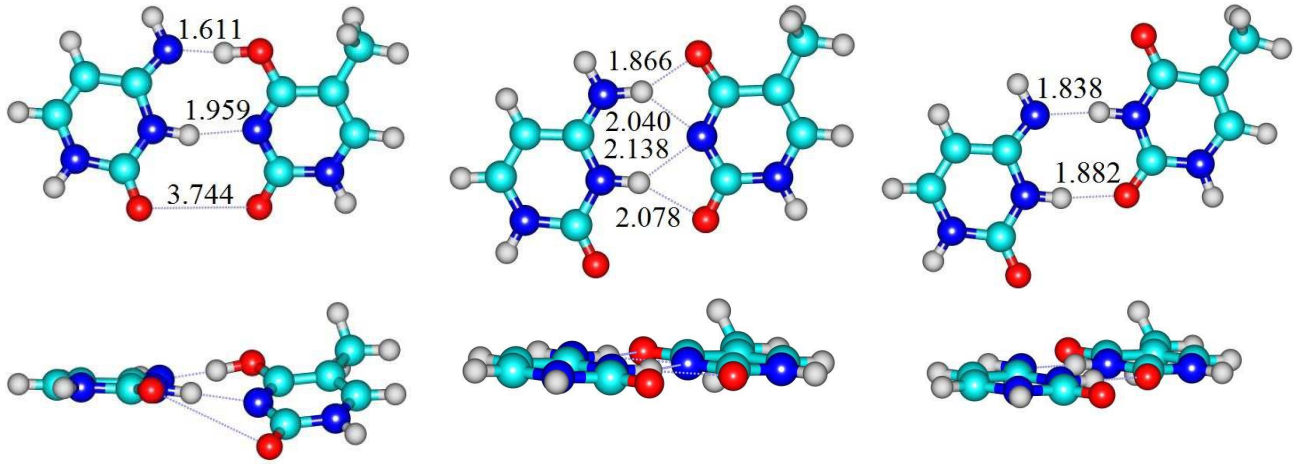
See designations in Table 1.
Energies of the H-bonds estimated by Nikolaienko–Bulavin–Hovorun formula [55] are denoted by double asterisk.
[£]Data are taken from the work [12].

(a)



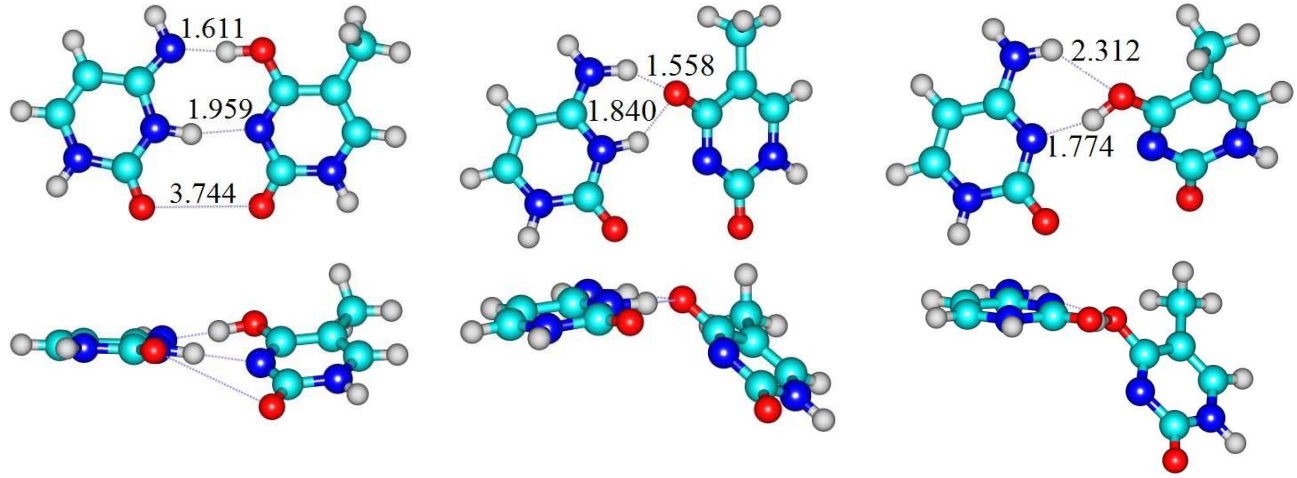
C·T(WC) | **$TS_{C\cdot T(WC)\leftrightarrow C^{*}\cdot T^{*}(WC)}$** ($\nu_i$=-701.8*i* cm^{-1}) | **C*·T*(WC)**

(b)



C*·T*(WC) | **$TS^{C+\cdot T-}_{C^{*}\cdot T^{*}(WC)\leftrightarrow C^{*}\cdot T\uparrow(w)}$** ($\nu_i$=-140.4*i* cm^{-1}) | **C*·$T_{\uparrow}$(w)**

(c)



C*·T*(WC) | **$TS^{C+\cdot T-}_{C^{*}\cdot T^{*}(WC)\leftrightarrow C\cdot T^{*}\downarrow(w)}$** ($\nu_i$=-55.8*i* cm^{-1}) | **C·T*$_{\downarrow}$(w)**

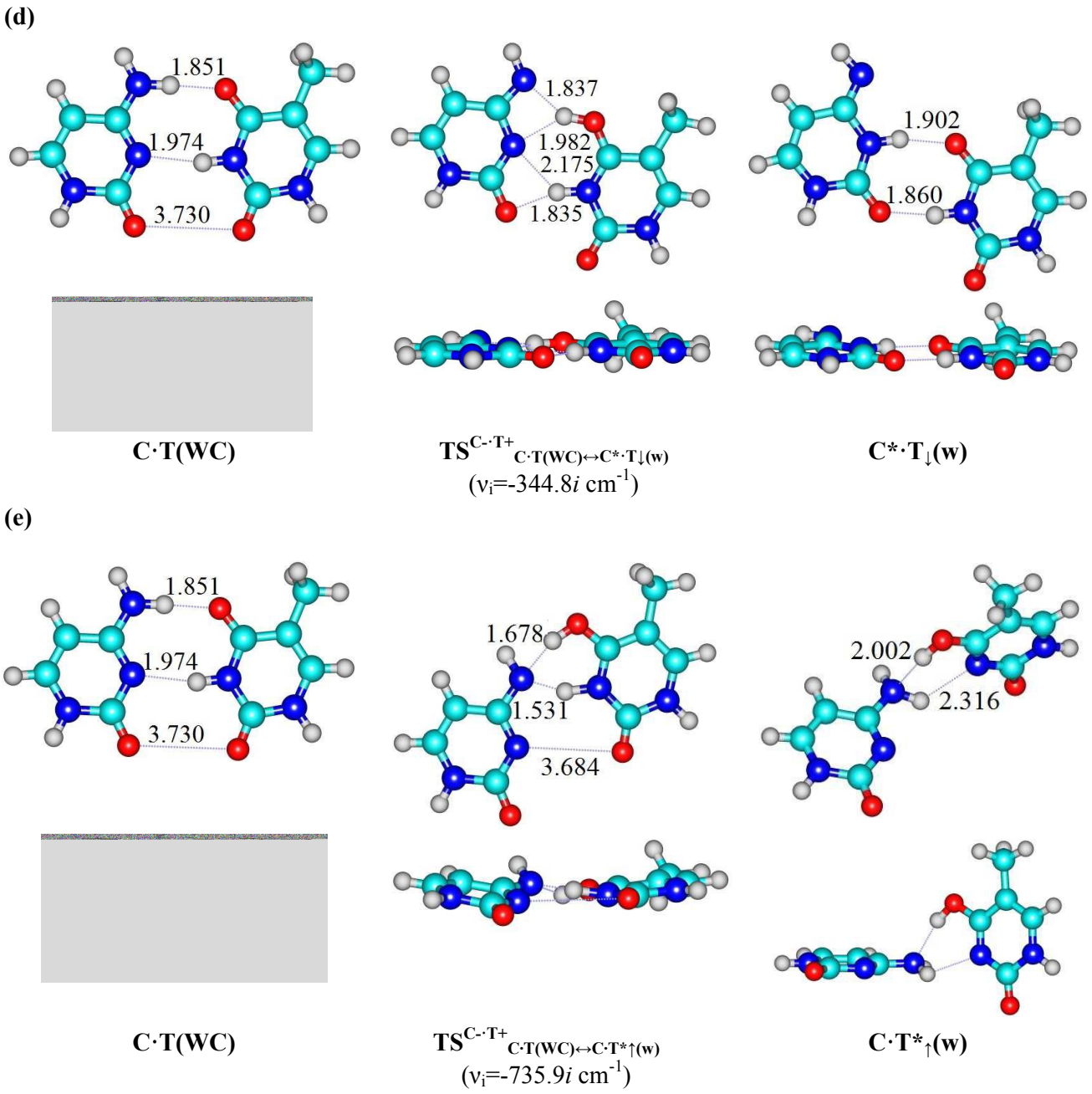
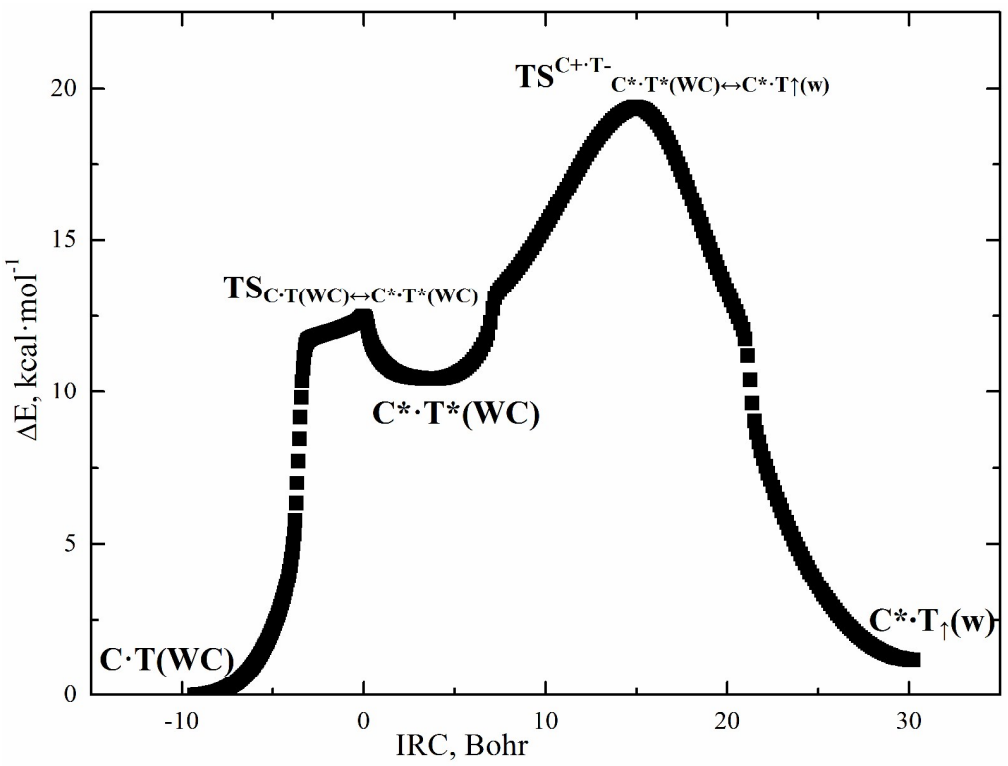
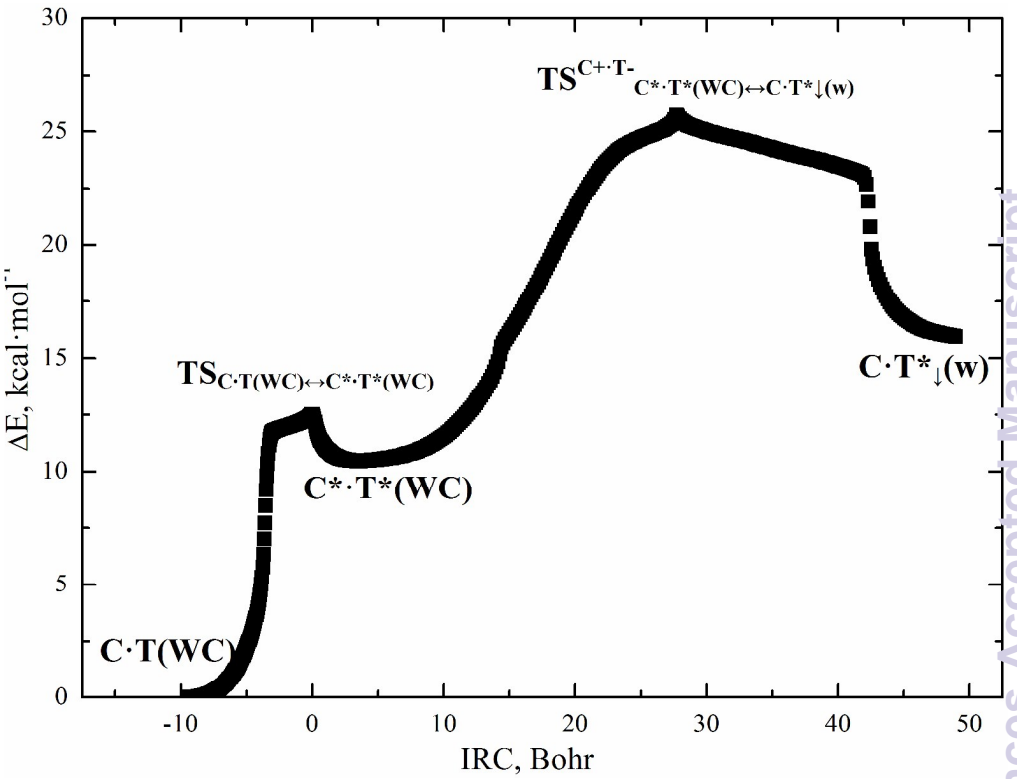


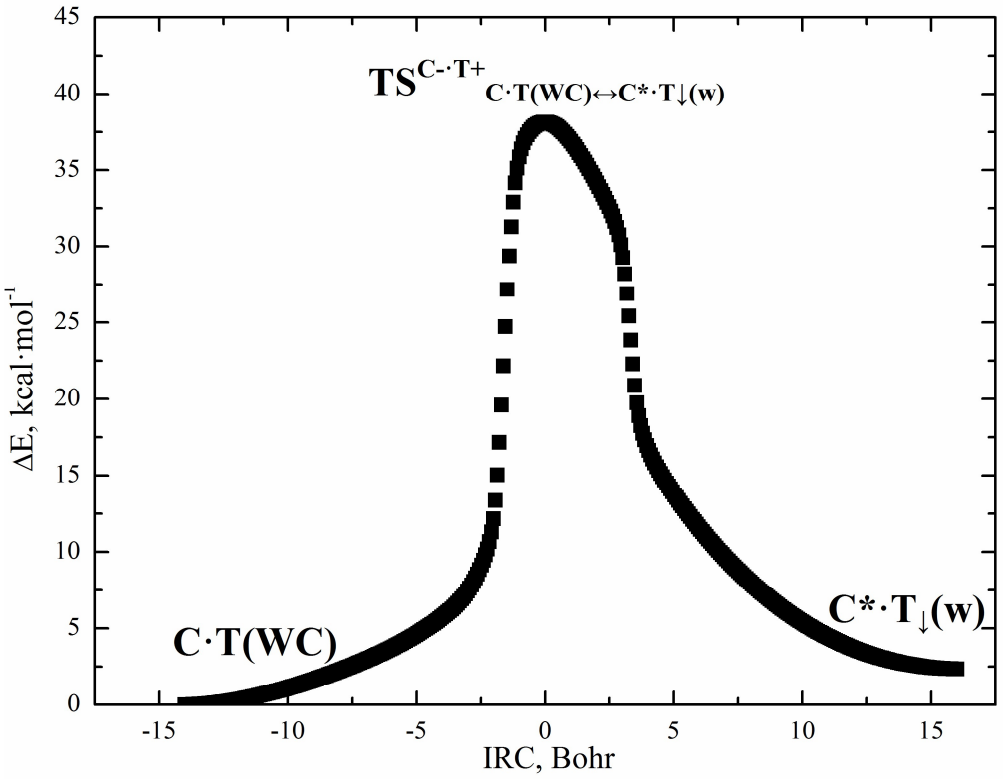
Fig. 4. Structures corresponding to the stationary points on the reaction pathways of the (**a**) C·T(WC)↔C*·T*(w), (**b**) C*·T*(WC)↔C*·T↑(w), (**c**) C*·T*(WC)↔C·T*↓(w), (**d**) C·T(WC)↔C*·T↓(w) and (**e**) C·T(WC)↔C·T*↑(w) conversions *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory. Dotted lines indicate AH···B H-bonds and attractive A···B van der Waals contacts (their lengths are presented in angstroms). Carbon atoms are in light-blue, nitrogen – in dark-blue, hydrogen – in grey and oxygen – in red. ν_i – imaginary frequency.



(a)



(b)



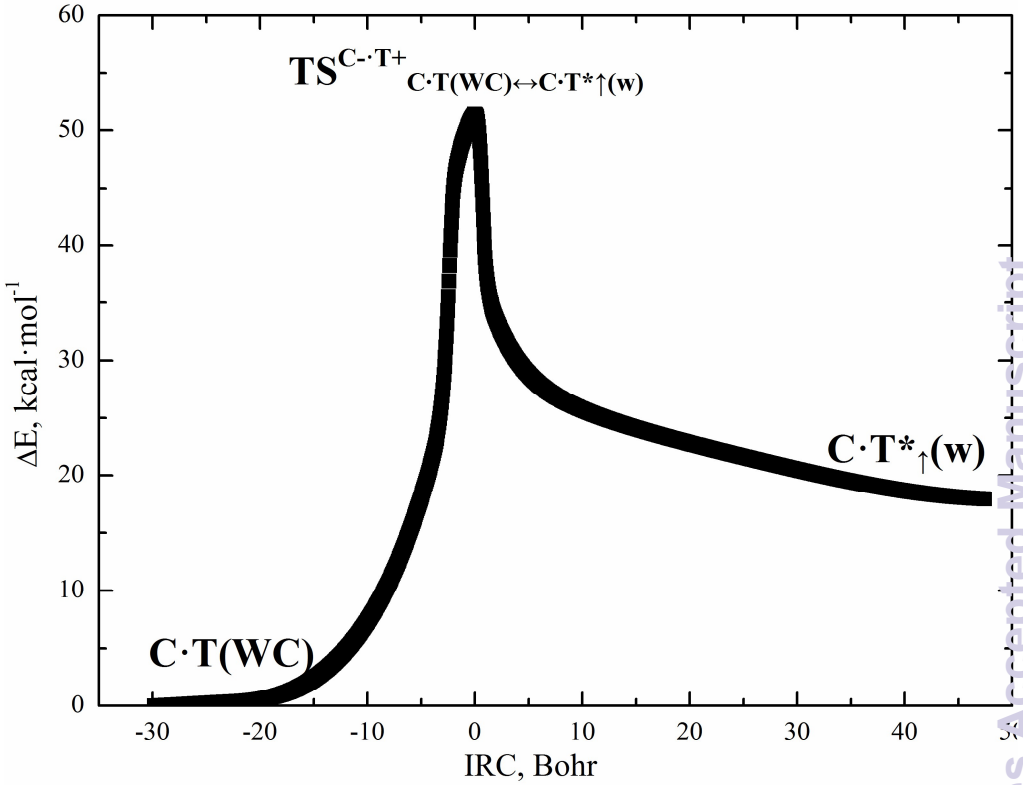


Fig. 5. Profiles of the relative electronic energy ΔE of the DNA base mispairs along the IRC of the (**a**) C·T(WC)↔C*·T$_{\uparrow}$(w), (**b**) C·T(WC)↔C·T*$_{\downarrow}$(w), (**c**) C·T(WC)↔C*·T$_{\downarrow}$(w) and (**d**) C·T(WC)↔C·T*$_{\uparrow}$(w) tautomerisations *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory.

Table 5. Energetic and kinetic characteristics of the C·T(WC)↔C*·T*(w), C*·T*(WC)↔C*·$T_{\uparrow}$(w), C*·T*(WC)↔C·T*$_{\downarrow}$(w), C·T(WC)↔C*·$T_{\downarrow}$(w) and C·T(WC)↔C·T*$_{\uparrow}$(w) tautomerisations *via* the sequential DPT of the C and T bases obtained at the different levels of theory for the geometry calculated at the B3LYP/6-311++G(d,p) level of theory (see also Fig. 4 and Table S1).

Level of theory	ΔG^a	ΔE^b	$\Delta\Delta G_{TS}{}^c$	$\Delta\Delta E_{TS}{}^d$	$\Delta\Delta G^e$	$\Delta\Delta E^f$	$\tau_{99.9\%}{}^g$
C·T(WC)↔C*·T*(WC)							
MP2/6-311++G(2df,pd)	8.97	8.81	9.08	10.91	0.11	2.10	$1.31\cdot10^{-13}$
MP2/6-311++G(3df,2pd)	9.24	9.08	9.54	11.37	0.30	2.29	$1.84\cdot10^{-13}$
MP2/cc-pVTZ	8.82	8.66	9.36	11.19	0.54	2.53	$2.70\cdot10^{-13}$
MP2/cc-pVQZ	9.15	8.99	9.55	11.38	0.40	2.39	$2.13\cdot10^{-13}$
C*·T*(WC)↔C*·$T_{\uparrow}$(w)							
MP2/6-311++G(2df,pd)	-8.31	-8.17	8.38	17.02	16.69	25.18	$1.52\cdot10^{-6}$
MP2/6-311++G(3df,2pd)	-8.68	-8.54	7.88	8.35	16.56	16.89	$6.61\cdot10^{-7}$
MP2/cc-pVTZ	-8.40	-8.25	8.54	9.01	16.94	17.27	$2.01\cdot10^{-6}$
MP2/cc-pVQZ	-8.59	-8.44	7.90	8.37	16.48	16.81	$6.76\cdot10^{-7}$
C*·T*(WC)↔C·T*$_{\downarrow}$(w)							
MP2/6-311++G(2df,pd)	2.87	5.53	17.74	16.58	14.87	11.05	$8.89\cdot10^{-2}$
MP2/6-311++G(3df,2pd)	2.86	5.53	17.44	16.28	14.57	10.75	$5.39\cdot10^{-2}$
MP2/cc-pVTZ	3.23	5.90	18.32	17.17	15.09	11.27	$1.30\cdot10^{-1}$
MP2/cc-pVQZ	2.92	5.58	17.49	16.33	14.57	10.75	$5.38\cdot10^{-2}$
C·T(WC)↔C*·$T_{\downarrow}$(w)							
MP2/6-311++G(2df,pd)	3.31	3.42	34.41	35.70	31.10	32.28	$6.49\cdot10^{10}$
MP2/6-311++G(3df,2pd)	2.97	3.09	34.23	35.52	31.26	32.43	$8.45\cdot10^{10}$
MP2/cc-pVTZ	3.09	3.20	34.66	35.95	31.57	32.75	$1.44\cdot10^{11}$
MP2/cc-pVQZ	3.10	3.22	34.40	35.70	31.30	32.48	$9.09\cdot10^{10}$
C·T(WC)↔C·T*$_{\uparrow}$(w)							
MP2/6-311++G(2df,pd)	15.60	16.97	46.51	47.72	30.91	30.75	$3.52\cdot10^{10}$
MP2/6-311++G(3df,2pd)	15.88	17.26	46.87	48.08	30.99	30.83	$4.00\cdot10^{10}$
MP2/cc-pVTZ	15.64	17.02	46.66	47.87	31.01	30.86	$4.20\cdot10^{10}$
MP2/cc-pVQZ	15.70	17.07	46.76	47.97	31.06	30.90	$4.55\cdot10^{10}$

[a]The Gibbs free energy of the product relatively the reactant of the tautomerisation reaction (T=298.15 K), kcal·mol^{-1}

[b]The electronic energy of the product relatively the reactant of the tautomerisation reaction, kcal·mol^{-1}

[c]The Gibbs free energy barrier for the forward reaction of tautomerisation, kcal·mol^{-1}

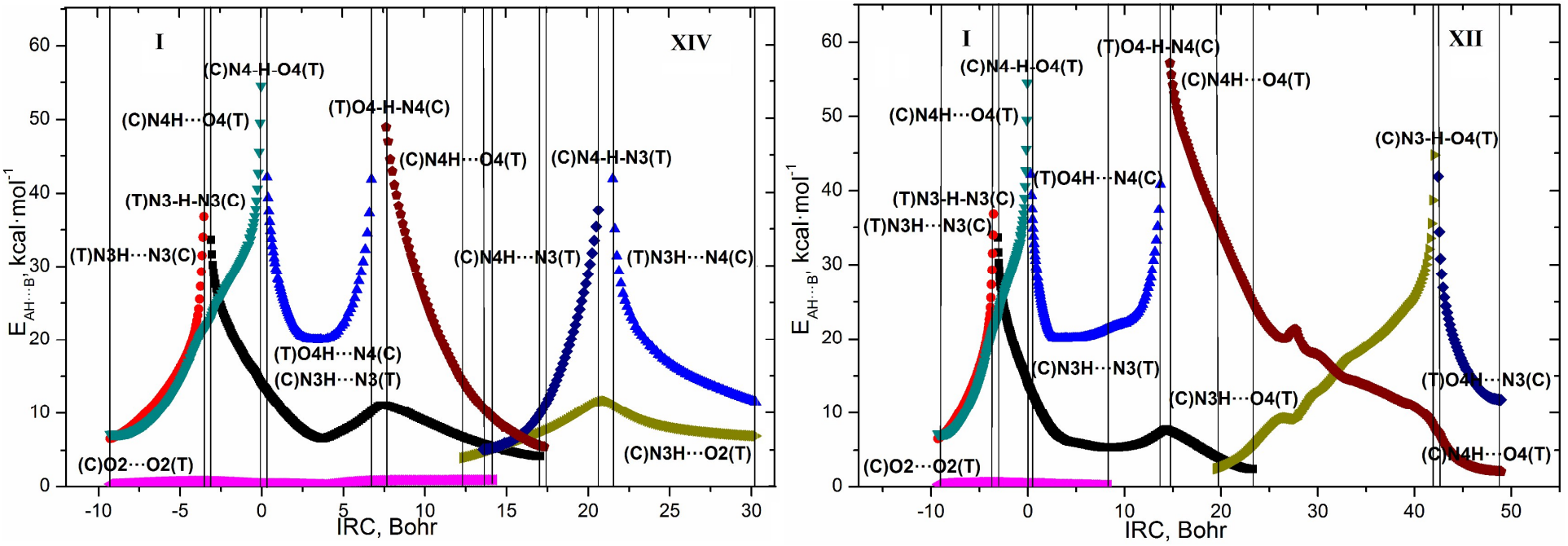
[d]The electronic energy barrier for the forward reaction of tautomerisation, kcal·mol^{-1}

[e]The Gibbs free energy barrier for the reverse reaction of tautomerisation, kcal·mol^{-1}

[f]The electronic energy barrier for the reverse reaction of tautomerisation, kcal·mol^{-1}

[g]The time necessary to reach 99.9%of the equilibrium concentration between the reactant and the product of the tautomerisation reaction, s

See also summary Table S1 for the Gibbs and electronic energies of the mispairs and TSs relatively the global minimum – the short C·T(WC) Watson-Crick DNA base mispair.



(a) **(b)**

Fig. 6. Exchange of the patterns of the intermolecular AH···B H-bonds (their energies $E_{AH\cdots B}$ are estimated by the EML formula at the (3,-1) BCPs) along the IRC of the (**a**) C·T(WC)↔C*·$T_{\uparrow}$(w) and (**b**) C·T(WC)↔C·T*$_{\downarrow}$(w) biologically important tautomerisations *via* the sequential DPT obtained at the B3LYP/6-311++G(d,p) level of theory (see Tables 4 and 6).

Table 6. Patterns of the intermolecular interactions including AH···B H-bonds and loosened A-H-B covalent bridges that sequentially replace each other along the IRC of the C·T(WC)↔C*·$T_{\uparrow}$(w) and C·T(WC)↔C·T*$_{\downarrow}$(w) biologically important tautomerisations *via* the sequential DPT and their ranges of the existence obtained at the B3LYP/6-311++G(d,p) level of theory (see Figs. 4-6).

Patterns	IRC range, Bohr	Intermolecular interactions, forming patterns
C·T(WC)↔C*·$T_{\uparrow}$(w)		
I	[-9.30÷-3.55)	(C)N4H···O4(T), (T)N3H···N3(C), (C)O2···O2(T)
II	[-3.55÷-3.12)	(C)N4H···O4(T), (T)N3-H-N3(C), (C)O2···O2(T)
III	[-3.12÷-0.05)	(C)N4H···O4(T), (C)N3H···N3(T), (C)O2···O2(T)
IV	[-0.05÷0.32)	(C)N4-H-O4(T), (C)N3H···N3(T), (C)O2···O2(T)
V	[0.32÷8.32)	(T)O4H···N4(C), (C)N3H···N3(T), (C)O2···O2(T)
VI	[8.32÷13.71)	(T)O4-H-N4(C), (C)N3H···N3(T), (C)O2···O2(T)
VII	[13.71÷14.70)	(C)N4H···O4(T), (C)N3H···N3(T), (C)O2···O2(T)
VIII	[14.70÷19.47)	(C)N4H···O4(T), (C)N3H···N3(T), (C)N3H···O2(T), (C)O2···O2(T)
IX	[19.47÷23.29)	(C)N4H···O4(T), (C)N4H···N3(T), (C)N3H···N3(T), (C)N3H···O2(T), (C)O2···O2(T)
X	[23.29÷42.02)	(C)N4H···O4(T), (C)N4H···N3(T), (C)N3H···N3(T), (C)N3H···O2(T)
XI	[42.02÷42.44)	(C)N4H···O4(T), (C)N4H···N3(T), (C)N3H···O2(T)
XII	[23.29÷42.02)	(C)N4H···N3(T), (C)N3H···O2(T)
XIII	[42.02÷42.44)	(C)N4-H-N3(T), (C)N3H···O2(T)
XIV	[42.44÷48.95]	(T)N3H···N4(C), (C)N3H···O2(T)
C·T(WC)↔C·T*$_{\downarrow}$(w)		
I	[-9.30÷-3.55)	(C)N4H···O4(T), (T)N3H···N3(C), (C)O2···O2(T)
II	[-3.55÷-3.12)	(C)N4H···O4(T), (T)N3-H-N3(C), (C)O2···O2(T)
III	[-3.12÷-0.05)	(C)N4H···O4(T), (C)N3H···N3(T), (C)O2···O2(T)
IV	[-0.05÷0.32)	(C)N4-H-O4(T), (C)N3H···N3(T), (C)O2···O2(T)
V	[0.32÷8.32)	(T)O4H···N4(C), (C)N3H···N3(T), (C)O2···O2(T)
VI	[8.32÷13.71)	(T)O4H···N4(C), (C)N3H···N3(T)
VII	[13.71÷14.70)	(T)O4-H-N4(C), (C)N3H···N3(T)
VIII	[14.70÷19.47)	(C)N4H···O4(T), (C)N3H···N3(T)
IX	[19.47÷23.29)	(C)N4H···O4(T), (C)N3H···O4(T), (C)N3H···N3(T)
X	[23.29÷42.02)	(C)N4H···O4(T), (C)N3H···O4(T)
XI	[42.02÷42.44)	(C)N4H···O4(T), (C)N3-H-O4(T)
XII	[42.44÷48.95]	(C)N4H···O4(T), (T)O4H···N3(C)